

GenreMbl:*
1: gb_ba:*
2: gb_hng:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vt:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mn:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vt:*
30: em_hng_hum:*
31: em_hng_inrv:*
32: em_hng_other:*
33: em_hng_mus:*
34: em_hng_pln:*
35: em_hng_rod:*
36: em_hng_mam:*
37: em_hng_vrrt:*
38: em_sy:*
39: em_hngo_hum:*
40: em_hngo_mus:*
41: em_hngo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	860.4	98.6	1574	9	BC012868	BC012868 Homo sapi
2	829	95.0	1628	9	AX014903	AX014903 Sequence
3	827	94.7	1558	9	BC001668	BC001668 Homo sapi
4	825.4	94.5	1424	9	BC011955	BC011955 Homo sapi
5	825.4	94.5	1457	9	BC011951	BC011951 Homo sapi
6	825.4	94.5	1509	9	BC007705	BC007705 Homo sapi
7	825.2	94.5	1544	6	AX410783	AX410783 Sequence
8	825.2	94.5	1544	9	HS078095	U78095 Homo sapien
9	809.4	92.7	1512	9	AB006534	AB006534 Homo sapi
10	805.6	92.3	1554	6	AX302243	AX302243 Sequence
11	805.6	92.3	1564	9	AF027205	AF027205 Homo sapi
12	621	71.1	16960	9	AY050668	AY050668 Homo sapi
13	621	71.1	16960	9	AC011479	AC011479 Homo sapi
14	595.2	68.2	64620	2	AC083794	AC083794 Homo sapi
15	394	45.1	759	6	AF069031	AF069031 Sequence
16	394	45.1	759	6	EL1900	EL1900 Human cDNA
17	394	45.1	759	9	194991	194991 H. sapiens (
18	337	38.6	349	9	H5X597	AX302973 Sequence
19	316.4	36.2	320	6	AX302873	AX337128 Sequence
20	312.2	35.8	331	6	AX409401	AX409401 Sequence
21	312.2	35.8	331	6	AX360381	AX360381 Sequence
22	273	31.3	285	6	AX610552	AX610552 Sequence
23	262	30.0	286	6	AX610552	BC003431 Mus muscu
24	255.2	29.2	1237	10	BC003431	AX261672 Sequence
25	253.8	29.1	1287	1	BC261672	BC026419 Mus muscu
26	252	28.9	1174	10	BC026419	AF093019 Mus muscu
27	252	28.9	1237	10	AF093019	AF093020 Mus muscu
28	252	28.9	1310	10	AF093020	AF093016 Mus muscu
29	252	28.9	1408	6	AF093016	AX303056 Sequence
30	238.4	27.3	241	6	AX303056	AX303056 Sequence
31	226.4	25.9	287	6	AX261607	AX261607 Sequence
32	200	22.9	200	6	AX379141	AX379141 Sequence
33	186.2	21.3	336	6	AX093200	AX093200 Sequence
34	167.4	19.2	276	6	AX366402	AX366402 Sequence
35	167	19.1	207	6	AX365889	AX365889 Sequence
36	167	19.1	224	6	AX261692	AX261692 Sequence
37	167	19.1	234	6	AX261412	AX261412 Sequence
38	167	19.1	234	6	AX261478	AX261478 Sequence
39	167	19.1	234	6	AX261740	AX261740 Sequence
40	167	19.1	234	6	AX379315	AX379315 Sequence
41	167	19.1	371	6	AX363658	AX363658 Sequence
42	165.4	18.9	234	6	AX261539	AX261539 Sequence
43	165.4	18.9	276	6	AX366401	AX366401 Sequence
44	150.6	16.3	302	6	AX261340	AX261340 Sequence
45	146.8	16.8	199	6	AX302805	AX302805 Sequence

ALIGNMENTS

RESULT 1	BC012868	1574 bp	mRNA	linear	PRI 22-AUG-2001
LOCUS	BC012868				
DEFINITION	BC012868				
ACCESSION	BC012868				
VERSION	BC012868.1				
KEYWORDS	MG.				
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 1574)				
TITLE	Strausberg,R.				
	Direct Submission				

JOURNAL

Submitted (20-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: gcgabs-remail.nih.gov
 Tissue Procurement: DCTD/DTF
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: villaloboscm.tmc.edu
 Villalobos, D.K., Luna, R.A., Hale, S.M., Hulik, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>
 Series: IPAK Plate: 21 Row: b Column: 23
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 2598967.

FEATURES

source

1..1574
 /organism="Homo sapiens"
 /db_xref="locusid:10653"
 /db_xref="taxon:9606"
 /clone="MGC:9154 IMAGE:3657277"
 /tissue_type="Ovary, adenocarcinoma"
 /clone_id="NIH_MGC_66"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 335..1093
 /codon_start=1
 /product="serine protease inhibitor, kunitz type, 2"
 /protein_id="AAH12668.1"
 /db_xref="gi:15277534"
 /translation="MAOLGLRRSRAPLALLGLSLVLAADRESIHDFCLVSKYVGRCAKSPRMVNVITDSCGLFVGGCGGNSNNITKEEDLKCAVTENATGLPATS RNADSSVPSAPRRQDSBDSHSDMNEYETANAYTPCRASFPKMTFEDRSCNN FIVGCRGNKNSYRSEACMLRCFRQGNPEPLVVLGLFVWLLPLGASNV YLIRARNRNOERALRTYVSSGDDKDLVAVYLV"

CDS

BASE COUNT 318 a 409 c 485 g 362 t
 ORIGIN

Query Match 98.6%; Score 860.4; DB 9; Length 1574;
 Best Local Similarity 99.8%; Pred. No. 4,4e-195;
 Matches 872; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CTCAGGATATTTCACTATGAAAGATCTGACGCGCAAGCAGTCACTGGGCGTTG 60
 DB 700 CTCAGGATATTTCACTATGAAAGATCTGACGCGCAAGCAGTCACTGGGCGTTG 759
 QY 61 CCGATGATCTTCCACGCTGTGACTTTGAGCTGAGAGAGAACTCTGCAATACTTCAT 120
 DB 760 CCGATGATCTTCCACGCTGTGACTTTGAGCTGAGAGAGAACTCTGCAATACTTCAT 819
 QY 121 CTATGAGAGCTGCGGGGCAATAGAACAGTACCGCTCTGAGAGAGCCCTGATGCTCG 180
 DB 820 CTATGAGAGCTGCGGGGCAATAGAACAGTACCGCTCTGAGAGAGCCCTGATGCTCG 879
 QY 181 CTGCTTCCGACAGAGAAATCTTCCCTGCGCCCTTGGCTCAAAGGAGTGGTCTGGG 240
 DB 880 CTGCTTCCGACAGAGAAATCTTCCCTGCGCCCTTGGCTCAAAGGAGTGGTCTGGG 939
 QY 241 GGGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
 DB 940 GGGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 999
 QY 301 GGTGCAAGAGAACCAAGAGAGTGGCTTGGCAAGCTGTGAGCTCCGAGATGACAA 360

DB 1000 GGTGCAAGAGAACCAAGAGAGTGGCTTGGCAAGCTGTGAGCTCCGAGATGACAA 1059
 QY 361 GGAGCAGCTGTGTAAGACATATGCTGTAACCGCGCTGTGCGCAAGAGACT-363 419
 DB 1060 GGAGCAGCTGTGTAAGACATATGCTGTAACCGCGCTGTGCGCAAGAGACTG363 1119
 QY 420 AAGGAGGAGAGACTATGTGTGAGCTTTTAAATAGAGGSAATTGACTCGATTGAGT 479
 DB 1120 AAGGAGGAGAGACTATGTGTGAGCTTTTAAATAGAGGSAATTGACTCGATTGAGT 1179
 QY 480 GATCATTAAGGAGTGAAGTGTCTCTGAGAGGTAGAGAGGCTGCTTCTGCTGAGCA 539
 DB 1180 GATCATTAAGGAGTGAAGTGTCTCTGAGAGGTAGAGAGGCTGCTTCTGCTGAGCA 1239
 QY 540 GGGATGGGTTTGTCTTGGAAATCTCTAGAGAGCTTCTCTGATATGCTGAGCTG 599
 DB 1240 GGGATGGGTTTGTCTTGGAAATCTCTAGAGAGCTTCTCTGATATGCTGAGCTG 1299
 QY 600 CAGCAGCCCCGAGTGTCTCTGCTGATGATGATTTCTTCTCAGGTAGATTCTT 659
 DB 1300 CAGCAGCCCCGAGTGTCTCTGCTGATGATGATTTCTTCTCAGGTAGATTCTT 1359
 QY 660 GCTTATGTAATTCATGCTGCTTCTTCTCATACAGAGTGAATGTGAAATGCTTCT 719
 DB 1360 GCTTATGTAATTCATGCTGCTTCTTCTCATACAGAGTGAATGTGAAATGCTTCT 1419
 QY 720 TTTGTTGCTGATTTATGCTTTTAACTATAAACAAAGTTTATTTAGATTCTG 779
 DB 1420 TTTGTTGCTGATTTATGCTTTTAACTATAAACAAAGTTTATTTAGATTCTG 1479
 QY 780 AAGAGAGAACTAAATGTCAGTTTAAATAAGAGGAGGCTTCCCTTAGATTAAT 839
 DB 1480 AAGAGAGAACTAAATGTCAGTTTAAATAAGAGGAGGCTTCCCTTAGATTAAT 1539
 QY 840 TTCAGATGCTCTCAAAAAAAAAAAAAAAAAA 873
 DB 1540 TTCAGATGCTCTCAAAAAAAAAAAAAAAAAA 1573

RESULT 2
 LOCUS AX014903 1628 bp DNA linear PAT 07-SEP-2000
 DEFINITION Sequence 102 from Patent WO9551040.
 ACCESSION AX014903
 VERSION AX014903.1 GI:10041170
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1628)
 AUTHORS Schmitt, A., Specht, T., Dahl, E., Hinzmann, B., Rosenthal, A. and Pilarsky, C.
 TITLE Human nucleic acid sequences from ovarian tumour tissue
 JOURNAL Patent: WO 95/3040-A-102.21-OCT-1999;
 SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
 BERND (DE); ROSENTHAL ANDRE (DE); METZGEN GES FUER GENOMFORSCHUN
 (DE); PILARSKY CHRISTIAN (DE)
 FEATURES
 source
 1..1628
 Location/Qualifiers
 BASE COUNT 329 a 433 c 498 g 368 t
 ORIGIN

Query Match 95.0%; Score 829; DB 6; Length 1628;
 Best Local Similarity 98.2%; Pred. No. 1.4e-187;
 Matches 849; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

QY 1 CTCAGGATATTTCACTATGAAAGATCTGACGCGCAAGCAGTCACTGGGCGTTG 60
 DB 764 CTCAGGATATTTCACTATGAAAGATCTGACGCGCAAGCAGTCACTGGGCGTTG 823

Db 1066 GAGAGAGCTGTGTAAGACATATGCTCTGTGACCGCCCTGTGCCAAGAGACTGGGG 1125

QY 420 AAGGAGGAGGAGACTATGTGTGAGCTTTTAAATAGAGGATTTGACTGGATTGAGT 479

Db 1126 AAGGAGGAGGAGACTATGTGTGAGCTTTTAAATAGAGGATTTGACTGGATTGAGT 1185

QY 480 GATCATTAAGGCTGAGGCTCTCTCTCTGGAGGATGAGAGGCTGCTCTCTGCTGCA 539

Db 1186 GATCATTAAGGCTGAGGCTCTCTCTCTGGAGGATGAGAGGCTGCTCTCTGCTGCA 1245

QY 540 GGGATGGGTTGCTTGGAAATCCTCTAGAGGCTCCCTCGGATGAGCTGAGTCTGG 599

Db 1246 GGGATGGGTTGCTTGGAAATCCTCTAGAGGCTCCCTCGGATGAGCTGAGTCTGG 1305

QY 600 CAGACAGCCCGAGTGTCTCTCTGCTGATGATTTCTTCTCCGAGGATGAGTCTCTT 659

Db 1306 CAGACAGCCCGAGTGTCTCTCTGCTGATGATTTCTTCTCCGAGGATGAGTCTCTT 1365

QY 660 GCTTATGTTGAATTCATGCTGCTCTCTCTCTCATCAAGAGATGATGGAATCGTTCT 719

Db 1366 GCTTATGTTGAATTCATGCTGCTCTCTCTCTCATCAAGAGATGATGGAATCGTTCT 1425

QY 720 TTGTTGCTGATTTAGTTTGGTTTAAACAAAGTTTAAATAGCAATCTG 779

Db 1426 TTGTTGCTGATTTAGTTTGGTTTAAACAAAGTTTAAATAGCAATCTG 1485

QY 780 AAGGAGGAGGATTAATGTACAGTTTAAATAGGAGGCTTCCCTTTAGAAATTA 838

Db 1486 AAGGAGGAGGATTAATGTACAGTTTAAATAGGAGGCTTCCCTTTAGAAATTA 1544

RESULT 4

BC011955 1424 bp mRNA linear PRI 02-AUG-2001

LOCUS

DEFINITION Homo sapiens, serine protease inhibitor, Kunitz type, 2, clone

ACCESSION BC011955

VERSION BC011955.1 GI:15080402

KEYWORDS

SOURCE

ORGANISM

Homo sapiens.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 (bases 1 to 1424)

AUTHORS

Strausberg, R.

TITLE

Direct Submission

JOURNAL

Submitted (30-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT

Contact: MGC help desk

Email: cgapbe-remail.nih.gov

Tissue Procurement: DCTD/DTF

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILN)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HOSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: villalob@bcm.tmc.edu

Villalob, D.K., Luna, R.A., Hale, S.M., Huylk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILN at: <http://image.llnl.gov>

Series: IRAX Plate: 21 Row: a Column: 22

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 2598967.

FEATURES

1. 1424

Location/Qualifiers

source

/organism="Homo sapiens"

Query Match 94.5%; Score 825.4; DB 9; Length 1424;

Best Local Similarity 99.8%; Pred. No. 1e-186;

Matches 837; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

BASE COUNT 292 a 367 c 426 g 339 t

ORIGIN

1 CTTCAAGGATATGTCATTAAGAAATCTGACCGCCCAAGCAAGTCACTGGCTTG 60

Db 573 CTTCAAGGATATGTCATTAAGAAATCTGACCGCCCAAGCAAGTCACTGGCTTG 632

QY 61 CCGTCAATCTCTCCACGCTGTAATTTGAGAGGAGAACTCTGCAATTAATTTGAT 120

Db 633 CCGTCAATCTCTCCACGCTGTAATTTGAGAGGAGAACTCTGCAATTAATTTGAT 682

QY 121 CTATGAGGCTGCTGGGCAATTAAGAAACAGTCACTGAGAGGCTTCATGCTCCG 180

Db 693 CTATGAGGCTGCTGGGCAATTAAGAAACAGTCACTGAGAGGCTTCATGCTCCG 752

QY 181 CTGCTCCGACAGAGAAATCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240

Db 753 CTGCTCCGACAGAGAAATCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 812

QY 241 GGGGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300

Db 813 GGGGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 872

QY 301 GGTGCAAGAGAGAACCAAGAGCTGCTGCGACCTGCTGCTGCTGCTGCTGCTG 360

Db 873 GGTGCAAGAGAGAACCAAGAGCTGCTGCGACCTGCTGCTGCTGCTGCTGCTG 932

QY 361 GGAGAGCTGCTGAGAACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 419

Db 933 GGAGAGCTGCTGAGAACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 992

QY 420 AAGGAGGAGGAGACTATGTGAGCTTTTAAATAGAGGATTTGACTGGATTGAGT 479

Db 993 AAGGAGGAGGAGACTATGTGAGCTTTTAAATAGAGGATTTGACTGGATTGAGT 1052

QY 480 GATCATTAAGGCTGAGGCTCTCTCTCTGGAGGATGAGAGGCTGCTCTGCTGCA 539

Db 1053 GATCATTAAGGCTGAGGCTCTCTCTCTGGAGGATGAGAGGCTGCTCTGCTGCA 1112

QY 540 GGGATGGGTTGCTTGGAAATCCTCTAGAGGCTCCCTCGGATGAGCTGAGTCTGG 599

Db 1113 GGGATGGGTTGCTTGGAAATCCTCTAGAGGCTCCCTCGGATGAGCTGAGTCTGG 1172

QY 600 CAGACAGCCCGAGTGTCTCTCTGCTGATGATTTCTTCTCCGAGGATGAGTCTCTT 659

Db 1173 CAGACAGCCCGAGTGTCTCTCTGCTGATGATTTCTTCTCCGAGGATGAGTCTCTT 1232

QY 660 GCTTATGTTGAATTCATGCTGCTCTCTCTCATCAAGAGATGATGGAATCGTTCT 719

Db 1233 GCTTATGTTGAATTCATGCTGCTCTCTCTCATCAAGAGATGATGGAATCGTTCT 1292

QY 720 TTGTTGCTGATTTAGTTTGGTTTAAACAAAGTTTAAATAGCAATCTG 779


```

Db 1293 TTGCTGCTGATTAATGTTTATTAAGTATTAACAAAGCTTTTATTAAGCATTCG 1352
Qy 780 AAGAGGAAAGTAAATGTAACAATTAAATAAAGGAGGCTTCCCTTTAGATATA 838
Db 1353 AAGAGGAAAGTAAATGTAACAATTAAATAAAGGAGGCTTCCCTTTAGATATA 1411

RESULT 5
BC011951 1457 bp mRNA linear PRI 02-AUG-2001
LOCUS Homo sapiens, serine protease inhibitor, Kunitz type, 2, clone
DEFINITION MGC:11711 IMAGE:3856970, mRNA, complete cds.
ACCESSION BC011951.1 GI:15080390
VERSION BC011951.1
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 1457)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (30-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
COMMENT NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: gcgpc@remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: villalob@bcm.tmc.edu.
Villalob, D.K., Luna, R.A., Hale, S.M., Huylk, S., Lu, X., Garcia,
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
Muzny, D.M., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LIML at: http://image.llnl.gov
Series: IRAK Plate: 21 Row: a Column: 14
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 2598967.
FEATURES
Source
1..1457
/organism="Homo sapiens"
/db_xref="locusid:10653"
/db_xref="taxon:9606"
/clone="MGC:11711 IMAGE:3856970"
/tissue_type="Ovary, adenocarcinoma"
/clone_id="NH_MGC_66"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
236..994
/codon_start=1
/product="serine protease inhibitor, Kunitz type, 2"
/protein_id="AAH11951.1"
/db_xref="GI:15080391"
/translation="MAQLGRLRSRAFLALLGLSLGLAADRSHIDFLVSKVY
GRCRASMRMYNVNTDSCQLFYGGCGGNSNVLTKBECLKCAIVENATGDLATS
RNAASISVPSAPRQDSSESSSMFYVEYCTANAYGPPRASFPPVYDPRNSCNV
FTYGGCRGNKNSITSEBACMLRCFRQENPPLPGSKVYVLLAGLFVWILFLGASVY
YLIRVARNRERARLTVSSGDKELVANYTVL"

BASE COUNT 304 a 374 c 439 g 340 t
ORIGIN
Query March 94.5% Score 825.4; DB 9; Length 1457;
Best Local Similarity 99.8%; Pred. No. 1e-186;
Matches 837; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

1 CTCCAGCGATATGTTGAACTATGAAATATCGACCGGCAAGCGAGTACGTCGGGCTTG 60
Db 601 CTCACGATATGTTGCACTATGAAATATCGACCGGCAAGCGAGTACGTCGGGCTTG 660
Qy 61 CCGTGATCTCTCCCAAGCTGTACTTTGACGTGAGAGAACTCCGCAATTAATTCAT 120
Db 661 CCGTGATCTCTCCCAAGCTGTACTTTGACGTGAGAGAACTCCGCAATTAATTCAT 720
Qy 121 CTATGAGAGCTGCGGAGCAATAAGAACAGCTACCGCTTGAGAGAGCTCGATGCTCGG 180
Db 721 CTATGAGAGCTGCGGAGCAATAAGAACAGCTACCGCTTGAGAGAGCTCGATGCTCGG 780
Qy 181 CTGCTTCCCGCAGCAGAGGAATCTCCCTCGCCCTTGAGCTCAAGGTGGTGGTCTGGC 240
Db 781 CTGCTTCCCGCAGCAGAGGAATCTCCCTCGCCCTTGAGCTCAAGGTGGTGGTCTGGC 840
Qy 241 GGGGCTGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
Db 841 GGGGCTGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
Qy 301 GATGACACGAGAGAACACAGAGCGTGCCTGCGACCGTCTGAGACTCCGAGATGACAA 360
Db 901 GATGACACGAGAGAACACAGAGCGTGCCTGCGACCGTCTGAGACTCCGAGATGACAA 960
Qy 361 GAGACGCTGTGTGAAGAACACATATGTCTGTGACCGCCCTGTGGCCAAAGGACT-GGG 419
Db 961 GAGACGCTGTGTGAAGAACACATATGTCTGTGACCGCCCTGTGGCCAAAGGACTGGG 1020
Qy 420 AAGGAGGAGGAGACATATGTGTGAGCTTTTAAATAGAGGATGATGATGATGATGATG 479
Db 1021 AAGGAGGAGGAGACATATGTGTGAGCTTTTAAATAGAGGATGATGATGATGATG 1080
Qy 480 GATCATTAAGGAGCTGAGGCTGTCTTCTGTGAGAGTGAAGAGGCTGTCTGTGAGCA 539
Db 1081 GATCATTAAGGAGCTGAGGCTGTCTTCTGTGAGAGTGAAGAGGCTGTCTGTGAGCA 1140
Qy 540 GGGATGGGTTTGTCTTGAAGAAATCTCTGAGAGGCTCTCTGCAATGGCTGCAATCTGG 599
Db 1141 GGGATGGGTTTGTCTTGAAGAAATCTCTGAGAGGCTCTCTGCAATGGCTGCAATCTGG 1200
Qy 600 CAGAGAGCCCGAGTGTCTTCTGCGCTGATGATGATGATGATGATGATGATGATGATG 659
Db 1201 CAGAGAGCCCGAGTGTCTTCTGCGCTGATGATGATGATGATGATGATGATGATGATG 1260
Qy 660 GCTTATGTAATTCATGCTCTTCTTCTCATCAGAGATGATGATGATGATGATGATGATG 719
Db 1261 GCTTATGTAATTCATGCTCTTCTTCTCATCAGAGATGATGATGATGATGATGATGATG 1320
Qy 720 TTGTTTGTCTGATTTATGTTTATTAAGTATTAACAAAGCTTTTATTAAGCATTCG 779
Db 1321 TTGTTTGTCTGATTTATGTTTATTAAGTATTAACAAAGCTTTTATTAAGCATTCG 1380
Qy 780 AAGAGGAAAGTAAATGTAACAATTAAATAAAGGAGGCTTCCCTTTAGATATA 838
Db 1381 AAGAGGAAAGTAAATGTAACAATTAAATAAAGGAGGCTTCCCTTTAGATATA 1439

RESULT 6
BC007705 1509 bp mRNA linear PRI 12-JUL-2001
LOCUS Homo sapiens, serine protease inhibitor, Kunitz type, 2, clone
DEFINITION MGC:11327 IMAGE:3952243, mRNA, complete cds.
ACCESSION BC007705
VERSION BC007705.1 GI:14043429
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 1509)
AUTHORS Strausberg, R.
TITLE Direct Submission

```

JOURNAL

Submitted (11-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs-remail.nih.gov

CDNA Library Preparation: DCTD/DTF
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Shcherbko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brickley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stantipop, S., Thomas, P.J., Tjongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: RAL Plate: 15 Row: n Column: 21
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

source

Location/Qualifiers
1..1509
/organism="Homo sapiens"
/db_xref="LocusID:10653"
/db_xref="taxon:9606"
/clone="MGC:11327 IMAGE:3952243"
/tissue_type="Ovary, adenocarcinoma"
/clone_id="NIH MGC-9"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
289..1047
/codon_start=1
/product="serine protease inhibitor, Kunitz type, 2"
/protein_id="AAH07705.1"
/db_xref="GI:14043430"
/translation="MAQLGLRRAFLALGSLLSGLVLAADRSIHDFVLSKVY
GCRASMRMYNYNTDSQCLFYGGCGSNMYLTKBECIKKCATATAGATGATS
RNAUUSVPSAPRRQDSSEDSMEYBYCTANNAVGPGRASFPVAPVDRNCNN
FLYGGCGNKNKSYSEACMLRCFRQENPPLPJSKVLHAGLFVWLILFLGASV
YLIRVARNOERALTIVWSGDDKEDLVKNYVL"

CDS

BASE COUNT 310 a 388 c 464 g 347 t
ORIGIN
Query Match 94.5%; Score 825.4; DB 9; Length 1509;
Best Local Similarity 99.8%; Pred. No. 1e-186;
Matches 837; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CTCACGCAATATGTTCACTATGAAGAATACATGACCGCCCAAGCAGTACTGGGCTTG 60
DB 654 CTCACGCAATATGTTCACTATGAAGAATACATGACCGCCCAAGCAGTACTGGGCTTG 713
QY 61 CCGTCATCTTCCACAGCGTGAATTTGACGTGAGAGAACTCTCGCAATTAATCTTAT 120
DB 714 CCGTCATCTTCCACAGCGTGAATTTGACGTGAGAGAACTCTCGCAATTAATCTTAT 773
QY 121 CTATGAGAGTGGCCGGGGCAATTAAGAACAGCTACCGCTCTGAGAGAGCTGATCTCG 180
DB 774 CTATGAGAGTGGCCGGGGCAATTAAGAACAGCTACCGCTCTGAGAGAGCTGATCTCG 833
QY 181 CTGCTTCGCGAGAGAGAACTCTCCCTGCGCTTGGCTCAAAAGTGTGTCTGGC 240
DB 834 CTGCTTCGCGAGAGAGAACTCTCCCTGCGCTTGGCTCAAAAGTGTGTCTGGC 893
QY 241 GGAGGCTGTGATGTGTGATCTCTTCTCTGGAGACCTCAATGGTCTAATCTGATCG 300

DB 894 GGAGGCTGTGATGTGTGATCTCTTCTCTGGAGACCTCAATGGTCTAATCTGATCG 953
QY 301 GGTGGACCGAGAGAACCAAGAGCGTGCCTCGGACCGTCTGAGAGCTCCGGAGATGACAA 360
DB 954 GGTGGACCGAGAGAACCAAGAGCGTGCCTCGGACCGTCTGAGAGCTCCGGAGATGACAA 1013
QY 361 GGAGGCTGTGAGAGAACCAATATGTCCTGGACCGCGCTGTGGCCAGAGAGACT-69G 419
DB 1014 GGAGGCTGTGAGAGAACCAATATGTCCTGGACCGCGCTGTGGCCAGAGAGACTG95G 1073
QY 420 AAGGAGGGGAGACATATGTGTGAGCTTTTAAATAGAGGATTTGACTCGATTGACT 479
DB 1074 AAGGAGGGGAGACATATGTGTGAGCTTTTAAATAGAGGATTTGACTCGATTGACT 1133
QY 480 GATCATTTAGGAGTGTGAGTCTTCTGTGGAGTAGAGAGCGTCTTCTGTGCTGGCA 539
DB 1134 GATCATTTAGGAGTGTGAGTCTTCTGTGGAGTAGAGAGCGTCTTCTGTGCTGGCA 1193
QY 540 GGGATGGCTTTTCTTGGAAATCTCTAGAGAGCGCTCTCTGCAATGGCGAGCTGG 599
DB 1194 GGGATGGCTTTTCTTGGAAATCTCTAGAGAGCGCTCTCTGCAATGGCGAGCTGG 1253
QY 600 CAGAGCCCGAGGTTTCTCTCGCTGATGATTTCTTCTCGAGTAGAGTTTCTT 659
DB 1254 CAGAGCCCGAGGTTTCTCTCGCTGATGATTTCTTCTCGAGTAGAGTTTCTT 1313
QY 660 GCTTATGTTGAATTCATGCTCTTCTTCTCATCACAGATGATTTGAATGCTTCT 719
DB 1314 GCTTATGTTGAATTCATGCTCTTCTTCTCATCACAGATGATTTGAATGCTTCT 1373
QY 720 TTGTGTTGTGTGATTTATGTTTATTAAGTATTAACAAAGCTTTTATATAGATCTG 779
DB 1374 TTGTGTTGTGTGATTTATGTTTATTAAGTATTAACAAAGCTTTTATATAGATCTG 1433
QY 780 AAAGAAGAAAGTAAATGTAACAATTAAATAAAGGCGCTTCCCTTACATATAA 838
DB 1434 AAAGAAGAAAGTAAATGTAACAATTAAATAAAGGCGCTTCCCTTACATATAA 1492
RESULT 7
AX410783 1544 bp DNA linear PAT 14-JUN-2002
LOCUS AX410783 Sequence 3430 from Patent WO0229103.
DEFINITION AX410783
ACCESSION AX410783.1 GI:21443468
VERSION
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Alvarez, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 3430 11-Apr-2002;
GENE LOGIC INC (US)
FEATURES
source
1..1544
/organism="Homo sapiens"
/db_xref="EMBL/GenBank Accession No. U78095"
/note="EMBL/GenBank Accession No. U78095"
BASE COUNT 316 a 394 c 473 g 360 t 1 others
ORIGIN
Query Match 94.5%; Score 825.2; DB 6; Length 1544;
Best Local Similarity 98.5%; Pred. No. 1.2e-186;
Matches 864; Conservative 0; Mismatches 9; Indels 4; Gaps 3;

QY 1 CTCACGCAATATGTTCACTATGAAGAATACATGACCGCCCAAGCAGTACTGGGCTTG 60
DB 666 CTCACGCAATATGTTCACTATGAAGAATACATGACCGCCCAAGCAGTACTGGGCTTG 725
QY 61 CCGTCATCTTCCACAGCGTGAATTTGACGTGAGAGAACTCTCGCAATTAATCTTAT 120

BASE COUNT 297 a 419 c 491 g 357 t
 ORIGIN /db_xref="taxon:9606"
 Query Match 92.3%; Score 805.6; DB 6; Length 1564;
 Best Local Similarity 99.3%; Pred. No. 5.6e-182;
 Matches 830; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY 1 CTCACGCGATATGTTCAACTATGAAAGTAATGACACCCCAACCGCATCTGAGGCTTG 60
 DB 730 CTCACGCGATATGTTCAACTATGAAAGTAATGACACCCCAACCGCATCTGAGGCTTG 789
 QY 61 CCGTGATGCTCCCGAGCGTGTGACTTTGACGCGAGAGAACTCCGCAATATCTCAT 120
 DB 790 CCGTGATGCTCCCGAGCGTGTGACTTTGACGCGAGAGAACTCCGCAATATCTCAT 849
 QY 121 CTATGAGAGCTCCCGGCGCAATAGAACAGCTACCGCTCTGAGAGGCTCGATGCTCCG 180
 DB 850 CTATGAGAGCTCCCGGCGCAATAGAACAGCTACCGCTCTGAGAGGCTCGATGCTCCG 909
 QY 181 CTGCTTCGCGCAGACAGAAATCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
 DB 910 CTGCTTCGCGCAGACAGAAATCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 969
 QY 241 GGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
 DB 970 GGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1029
 QY 301 GGTGGCAGGAGAACTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
 DB 1030 GGTGGCAGGAGAACTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1089
 QY 361 GGAACAGCTGCTGAGAAACACATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 419
 DB 1090 GGAACAGCTGCTGAGAAACACATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1149
 QY 420 AAGGAGGAGGAGACTATGCTGAGCTTTTAAATAGAGGAGTGAATCGGATTTGAGT 479
 DB 1150 AAGGAGGAGGAGACTATGCTGAGCTTTTAAATAGAGGAGTGAATCGGATTTGAGT 1209
 QY 480 GATCATTAAGGGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 539
 DB 1210 GATCATTAAGGGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1269
 QY 540 GGAATGGGTTGCTTTGAAATCTCTAGAGGCTCTCTCTGCAATGAGCTGAGTCTGG 599
 DB 1270 GGAATGGGTTGCTTTGAAATCTCTAGAGGCTCTCTCTGCAATGAGCTGAGTCTGG 1329
 QY 600 CAGACGCGCGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 659
 DB 1330 CAGACGCGCGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1389
 QY 660 GCTATATGTAATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 719
 DB 1390 GCTATATGTAATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1449
 QY 720 TTTGTTGCTGATTTATGTTTAAATAGAAAGTTTAAATAGAAATTTTAAATAGAA 779
 DB 1450 TTTGTTGCTGATTTATGTTTAAATAGAAAGTTTAAATAGAAATTTTAAATAGAA 1508
 QY 780 AAAGAGGAAAGTAAATGTACAAAGTTTAAATAGAAAGGCTCTCCCTTTAGAAAT 835
 DB 1509 AAAGAGGAAAGTAAATGTACAAAGTTTAAATAGAAAGGCTCTCCCTTTAGAAAT 1564

RESULT 11
 AF027205 1564 bp mRNA linear PRI 09-NOV-1997
 LOCUS Homo sapiens Kunitz-type protease inhibitor (kop) mRNA, complete
 DEFINITION cds.
 ACCESSION AF027205
 VERSION AF027205.1 GI:2598967
 KEYWORDS

SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1564)
 AUTHORS Mueller-Pillasch, F., Wallrapp, C., Bartels, K., Varga, G., Friess, H.,
 Buechler, M., Adler, G. and Gress, T.M.
 TITLE Cloning of a new Kunitz-type protease inhibitor with a putative
 transmembrane domain overexpressed in pancreatic cancer
 JOURNAL Biochim. Biophys. Acta (1997) In press
 REFERENCE 2 (bases 1 to 1564)
 AUTHORS Mueller-Pillasch, F., Wallrapp, C., Bartels, K., Adler, G. and
 Gress, T.M.
 TITLE Direct Submission
 JOURNAL Submitted (29-SEP-1997) Medizinische Klinik, Internal Medicine I,
 Robert-Koch-Str 8, Ulm 89081, Germany
 FEATURES
 source
 1. 1564
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="19"
 /map="19q13.1"
 /cell_line="Pau 8988"
 /tissue_type="pancreatic cancer"
 1. 1564
 /gene="kop"
 /protein_id="AA84031.1"
 /translation="MAHLGLRRSFAFLALIGSLIVLADREKSIHDEKYSKY
 GRCAAMPKMYVNTDSCOLFVYGGCDGNMNYLTKBECLKKCAIVENATGDLATIS
 RNADSSVSPAPRRDSEDSHSDMFNEYECYTAANAATPCRAFPFMYEDVRNCCN
 RYVGGCGKNYSRSEACMLRCFQDENPLPLGSRVYVLAGLFWVILILPLASMY
 YLIRVARNRERALRTVYSSGDKKQLKNTYV"
 BASE COUNT 297 a 419 c 491 g 357 t
 ORIGIN

Query Match 92.3%; Score 805.6; DB 9; Length 1564;
 Best Local Similarity 99.3%; Pred. No. 5.6e-182;
 Matches 830; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY 1 CTCACGCGATATGTTCAACTATGAAAGTAATGACACCCCAACCGCATCTGAGGCTTG 60
 DB 730 CTCACGCGATATGTTCAACTATGAAAGTAATGACACCCCAACCGCATCTGAGGCTTG 789
 QY 61 CCGTGATGCTCCCGAGCGTGTGACTTTGACGCGAGAGAACTCCGCAATATCTCAT 120
 DB 790 CCGTGATGCTCCCGAGCGTGTGACTTTGACGCGAGAGAACTCCGCAATATCTCAT 849
 QY 121 CTATGAGAGCTCCCGGCGCAATAGAACAGCTACCGCTCTGAGAGGCTCGATGCTCCG 180
 DB 850 CTATGAGAGCTCCCGGCGCAATAGAACAGCTACCGCTCTGAGAGGCTCGATGCTCCG 909
 QY 181 CTGCTTCGCGCAGACAGAAATCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
 DB 910 CTGCTTCGCGCAGACAGAAATCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 969
 QY 241 GGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
 DB 970 GGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1029
 QY 301 GGTGGCAGGAGAACTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
 DB 1030 GGTGGCAGGAGAACTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1089
 QY 361 GGAACAGCTGCTGAGAAACACATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 419

LOCUS AC011479 169660 bp DNA linear PRI 14-JUL-2002
 DEFINITION Homo sapiens chromosome 19 clone CTB-102L5, complete sequence.
 AC011479
 AC011479.6 GI:9954635
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 169660)
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
 TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 169660)
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 REFERENCE 3 (bases 1 to 169660)
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
 TITLE Direct Submission
 JOURNAL Submitted (31-AUG-2000) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 REFERENCE 4 (bases 1 to 169660)
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
 TITLE Direct Submission
 JOURNAL Submitted (14-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 On Aug 31, 2000 this sequence version replaced gi:7711511.
 Draft Sequence Produced by DOE Joint Genome Institute
 COMMENT www.jgi.doe.gov
 Finishing Completed at Stanford Human Genome Center
 www-shgc.stanford.edu
 Quality: Phrap Quality >=40 99.8% of Sequence;
 Estimated Total Number of Errors is 1.2.
 STS Content:
 SHGC-141352 G63443
 Location/Qualifiers
 source 1..169660
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="19"
 /clone="CTB-102L5"
 BASE COUNT 39214 a 45115 c 44847 g 40484 t
 ORIGIN
 Query Match 71.1%; Score 621; DB 9; Length 169660;
 Best Local Similarity 94.1%; Pred. No. 1.8e-137;
 Matches 656; Conservative 0; Mismatches 40; Indels 1; Gaps 1;
 Db 160 TGAGGAGGCTCGATCTCCGCTCTCCGACAGAGAGAACTCCCTCCCTCCCTGG 219
 Db 27605 TGGTGAAGCGCACTCGGCTGCACTCCCTTCCTGAGCCGCTCTCTC 27546
 Qy 220 CTGAAGGCTGGTCTTCTGGCGGGGCTTCTGATGCTGATCCCTTCTCTGGAGC 279
 Db 27545 GTCCAGTGTCTGCTGCTGCGGGGCTGCTGATGCTGATCCCTTCTCTGGAGC 27486
 Qy 280 CTCGATGCTCTCTGATCCCGGTGGGAGGAGACGAGAGCGTGCCTGGGACCCG 339
 Db 27485 CTCGATGCTCTCTGATCCCGGTGGGAGGAGACGAGAGCGTGCCTGGGACCCG 27426
 Qy 340 CTCGAGCTCCGAGATGACAGAGAGCACTGTGAAGAACATATGTCTGTGACCGCC 399
 Db 27425 CTCGAGCTCCGAGATGACAGAGAGCACTGTGAAGAACATATGTCTGTGACCGCC 27366
 Qy 400 CTGTCCGCAAGAGACT GGGAGGGAGGGGAGACATATGTGTGACCTTTTAAATAGA 458
 Db 27365 CTGTCCGCAAGAGACTGTGGGAGGGGAGAGGAGACATATGTGTGACCTTTTAAATAGA 27306
 Qy 459 GGGATTAACCTCGATTGAGTATGATATAGGAGCTGTGTTCTCTGGAGAGTAGA 518
 Db 27305 GGGATTAACCTCGATTGAGTATGATATAGGAGCTGTGTTCTCTGGAGAGTAGA 27246

Qy 519 CGAGTCTCTCTGCTGCTGCGAGGAGTGGTTCCTTGGAAATCTCTAGAGCTCTC 578
 Db 27245 CGAGTCTCTCTGCTGCTGCGAGGAGTGGTTCCTTGGAAATCTCTAGAGCTCTC 27186
 Qy 579 CTGGATGGCTCTGAGCTGCGAGAGCCCGAGTGTTCCTCGCTGATTCATTTCT 638
 Db 27185 CTGGATGGCTCTGAGCTGCGAGAGCCCGAGTGTTCCTCGCTGATTCATTTCT 27126
 Qy 639 CTTCCAGGAGAGTTCCTTGTGCTATGTTGAATCCATTCCTCTTCTCATCAGAGA 698
 Db 27125 CTTCCAGGAGAGTTCCTTGTGCTATGTTGAATCCATTCCTCTTCTCATCAGAGA 27066
 Qy 699 AGTATGTTGAATCTGCTTCTTGTGCTGATTAATGATTTTAAATAGAA 758
 Db 27065 AGTATGTTGAATCTGCTTCTTGTGCTGATTAATGATTTTAAATAGAA 27006
 Qy 759 AAGTTTATTTAGCATTTCTGAAAGAGAAATGTAATGTAAGTTTAAATAGAGG 818
 Db 27005 AAGTTTATTTAGCATTTCTGAAAGAGAAATGTAATGTAAGTTTAAATAGAGG 26946
 Qy 819 GCCTCCCTTTAGAAATTAATTCAGCATGCTTC 855
 Db 26945 GCCTCCCTTTAGAAATTAATTCAGCATGCTTC 26909
 RESULT 14
 AC083794/c 64620 bp DNA linear HTG 30-SEP-2000
 LOCUS AC083794/c
 DEFINITION Homo sapiens chromosome 4 clone RP11-140B3 map 4, LOW-PASS SEQUENCE
 SAMPLING.
 AC083794 64620 bp DNA linear HTG 30-SEP-2000
 VERSION AC083794.1 GI:10440701
 KEYWORDS HTG; HTGS; PHASED.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 64620)
 AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
 JOURNAL Homo sapiens chromosome 4, clone RP11-140B3
 REFERENCE 2 (bases 1 to 64620)
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S., Barna, N., Bastien, V., Beda, F., Boguslavsky, L., Boukhalter, B., Brown, A., Burkett, G., Campolano, A., Castle, A., Choquel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardy, S., Glend, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Harford, A., Horton, L., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., LaRoque, K., Lamazares, R., Landers, T., Lenoczky, J., Levine, R., Lieu, C., Liu, G., Macdonald, P., Margulis, N., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrum, J., Menes, L., Mhova, T., Meng, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riback, M., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Sougnuez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodores, J., Tirrell, A., Travers, M., Triggillo, J., Vasilev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zaitoun, J., Zimmerman, A. and Zody, M.
 Direct Submission
 Submitted (30-SEP-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/kw/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center Project name: L1254
Center clone name: 140_E_3

* NOTE: This record contains 78 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

* 1
* 715 814: contig of 714 bp in length
* 815 1522: contig of 708 bp in length
* 1523 1622: gap of 100 bp
* 1623 2347: contig of 725 bp in length
* 2348 2447: gap of 100 bp
* 2448 3182: contig of 735 bp in length
* 3183 3282: gap of 100 bp
* 3283 4020: contig of 738 bp in length
* 4021 4120: gap of 100 bp
* 4121 4848: contig of 728 bp in length
* 4849 4948: gap of 100 bp
* 4949 5677: contig of 729 bp in length
* 5678 5777: gap of 100 bp
* 5778 6523: contig of 746 bp in length
* 6524 6623: gap of 100 bp
* 6624 7362: contig of 739 bp in length
* 7363 7462: gap of 100 bp
* 7463 8180: contig of 728 bp in length
* 8191 8290: gap of 100 bp
* 8291 9000: contig of 710 bp in length
* 9001 9100: gap of 100 bp
* 9101 9825: contig of 725 bp in length
* 9826 9925: gap of 100 bp
* 9926 10646: contig of 721 bp in length
* 10647 10746: gap of 100 bp
* 10747 11461: contig of 715 bp in length
* 11462 11561: gap of 100 bp
* 11562 12292: contig of 731 bp in length
* 12293 12392: gap of 100 bp
* 12393 13135: contig of 743 bp in length
* 13136 13235: gap of 100 bp
* 13236 13959: contig of 724 bp in length
* 13960 14059: gap of 100 bp
* 14060 14795: contig of 736 bp in length
* 14796 14895: gap of 100 bp
* 14896 15642: contig of 747 bp in length
* 15643 15742: gap of 100 bp
* 15743 16462: contig of 720 bp in length
* 16463 16562: gap of 100 bp
* 16563 17306: contig of 744 bp in length
* 17307 17406: gap of 100 bp
* 17407 18117: contig of 711 bp in length
* 18118 18217: gap of 100 bp
* 18218 18945: contig of 728 bp in length
* 18946 19045: gap of 100 bp
* 19046 19761: contig of 736 bp in length
* 19762 19861: gap of 100 bp
* 19862 20605: contig of 724 bp in length
* 20606 20705: gap of 100 bp
* 20706 21446: contig of 741 bp in length
* 21447 21546: gap of 100 bp
* 21547 22282: contig of 736 bp in length
* 22283 22382: gap of 100 bp
* 22383 23125: contig of 743 bp in length
* 23126 23225: gap of 100 bp
* 23226 23947: contig of 722 bp in length

* 23948 24047: gap of 100 bp
* 24048 24776: contig of 729 bp in length
* 24777 24876: gap of 100 bp
* 24877 25594: contig of 718 bp in length
* 25595 25694: gap of 100 bp
* 25695 26441: contig of 747 bp in length
* 26442 26541: gap of 100 bp
* 26542 27283: contig of 742 bp in length
* 27284 27383: gap of 100 bp
* 27384 28121: contig of 738 bp in length
* 28122 28221: gap of 100 bp
* 28222 28957: contig of 736 bp in length
* 28958 29057: gap of 100 bp
* 29058 29808: contig of 751 bp in length
* 29809 29908: gap of 100 bp
* 29909 30651: contig of 743 bp in length
* 30652 30751: gap of 100 bp
* 30752 31508: contig of 757 bp in length
* 31509 31608: gap of 100 bp
* 31609 32337: contig of 729 bp in length
* 32338 32437: gap of 100 bp
* 32438 33175: contig of 738 bp in length
* 33176 33275: gap of 100 bp
* 33276 34002: contig of 726 bp in length
* 34002 34101: gap of 100 bp
* 34102 34799: contig of 698 bp in length
* 34800 34899: gap of 100 bp
* 34900 35623: contig of 724 bp in length
* 35624 35723: gap of 100 bp
* 35724 36453: contig of 730 bp in length
* 36454 36553: gap of 100 bp
* 36554 37271: contig of 718 bp in length
* 37272 37371: gap of 100 bp
* 37372 38098: contig of 727 bp in length
* 38099 38198: gap of 100 bp
* 38199 38906: contig of 708 bp in length
* 38907 39006: gap of 100 bp
* 39007 39719: contig of 713 bp in length
* 39720 39819: gap of 100 bp
* 39820 40520: contig of 701 bp in length
* 40521 40620: gap of 100 bp
* 40621 41363: contig of 743 bp in length
* 41364 41463: gap of 100 bp
* 41464 42184: contig of 721 bp in length
* 42185 42284: gap of 100 bp
* 42285 43030: contig of 746 bp in length
* 43031 43130: gap of 100 bp
* 43131 43865: contig of 735 bp in length
* 43866 43965: gap of 100 bp
* 43966 44706: contig of 741 bp in length
* 44707 44806: gap of 100 bp
* 44807 45521: contig of 715 bp in length
* 45522 45621: gap of 100 bp
* 45622 46341: contig of 720 bp in length
* 46342 46441: gap of 100 bp
* 46442 47166: contig of 725 bp in length
* 47167 47266: gap of 100 bp
* 47267 47981: contig of 715 bp in length
* 47982 48081: gap of 100 bp
* 48082 48817: contig of 736 bp in length
* 48818 48917: gap of 100 bp
* 48918 49635: contig of 718 bp in length
* 49636 49735: gap of 100 bp
* 49736 50462: contig of 727 bp in length
* 50463 50562: gap of 100 bp
* 50563 51287: contig of 725 bp in length
* 51288 51387: gap of 100 bp
* 51388 52107: contig of 720 bp in length
* 52108 52207: gap of 100 bp
* 52208 52943: contig of 736 bp in length
* 52944 53043: gap of 100 bp
* 53044 53774: contig of 731 bp in length
* 53775 53874: gap of 100 bp


```
*
538975 54611: config of 737 bp in length
54612 54711: gap of 100 bp
54712 55429: config of 718 bp in length
55430 55529: gap of 100 bp
55530 56244: config of 715 bp in length
56245 56344: gap of 100 bp
56345 57099: config of 754 bp in length
57099 57190: gap of 100 bp
57191 57904: config of 714 bp in length
```

Query Match	68.2%;	Score 595.2;	DB 2;	Length 64620;
Best Local Similarity	96.1%;	Pred. No. 2.2e-131;		
Matches 620;	Conservative	0;	Mismatches 24;	Indels 1;
				Gaps 1

QY	202	TTCTCCCTGCGCCCTTGGATGCTAAAGGAGGAGGCTCTGGAGGAGCTCTTCGATAGGTT	261
Db	3359	TCTTAGGCCCCCTCTGATCTTAGGAGGCTCTGGAGGAGCTCTTCGATAGGTT	3900
QY	262	GATCCTCTCTCTGGAGGCTCCATGAGCTACACCTGATCCGGGTGGACGAGAACGACGAG	321
Db	3899	GATCCTCTCTCTGGAGGCTCCATGAGCTACACCTGATCCGGGTGGACGAGAACGACGAG	3840
QY	322	GCGTCCCTCGGACCGCTGTGAGACTTCGGAATGACAAAGACAGCTGTGGAAACAC	381
Db	3839	GCGTCCCTCGGACCGCTGTGAGACTTCGGAATGACAAAGACAGCTGTGGAAACAC	3780
QY	382	ATATGCTCCTGACCCGCTCTGCCAAGAGACT - GGAAGGAGGAGGAGACTATGTT	440
Db	3779	ATATGCTCCTGACCCGCTCTGCCAAGAGACT - GGAAGGAGGAGGAGACTATGTT	3720
QY	441	GAGCTTTTAAATAGAGGATTAAGTGGATTTAGGCTGAGCTCG	500
Db	3719	GAGCTTTTAAATAGAGGATTAAGTGGATTTAGGCTGAGCTCG	3660
QY	501	TTTCTCGAGAGGTAGAGGCTCTCTCGATCTGGACAGGATGGTTTCTTGGAA	560
Db	3659	TTTCTCGAGAGGTAGAGGCTCTCTCGATCTGGACAGGATGGTTTCTTGGAA	3600
QY	561	TCCTCTAGAGGCTCTCTCGATCTGGACCTGAGCTGAGAGAGCCCGATGTTTCC	620
Db	3599	TCCTCTAGAGGCTCTCTCGATCTGGACCTGAGCTGAGAGAGCCCGATGTTTCC	3540
QY	621	TGCGTGAATCATTTCTTTCCTCCAGGTAGAGTTTCTTTCCTTATGTTTCCATTC	680
Db	3539	TGCGTGAATCATTTCTTTCCTCCAGGTAGAGTTTCTTTCCTTATGTTTCCATTC	3480
QY	681	CTCTTTCTCATCAGAGAGTGGATCGTATCTTTTGTGTCTATTTAGT	740
Db	3479	CTCTTTCTCATCAGAGAGTGGATCGTATCTTTTGTGTCTATTTAGT	3420
QY	741	TTTTTAAATAAACAAAGTTTTTATTAGCATCTGAAGACGAAAGTAAATGTA	800
Db	3419	TTTTTAAATAAACAAAGTTTTTATTAGCATCTGAAGACGAAAGTAAATGTA	3360
QY	801	CAAGTTAATAAAAGGAGCTTCCCTTAAATTAATTTGAC	845
Db	3359	CAAGTTAATAAAAGGAGCTTCCCTTAAATTAATTTGAC	3315

RESULT 15	AR069031	AR069031	759 bp	DNA	linear	PAT 29-SEP-1996
LOCUS	AR069031	Sequence 4 from patent US 5854396.				
DEFINITION	AR069031					
ACCESSION	AR069031					
VERSION	AR069031.1	GI:6001238				
KEYWORDS						
SOURCE	Unknown.					
ORGANISM	Unknown.					
	Unclassified.					

REFERENCE
1 (bases 1 to 759)
AUTHORS
Shimomura, T., Kawaguchi, T. and Kilmura, N.
TITLE
Protein, DNA coding for same and method of producing the protein
JOURNAL
Patent: US 5854396-A 4 29-DEC-1998;

FEATURES	Location/Qualifiers
source	1. .759
	/organism="unknown"
BASE COUNT	152 a 213 c 236 g 158 t
ORIGIN	

Query Match 45.1%; Score 394; DB 6; Length 759;
Best Local Similarity 100.0%; Pred. No. 1.5e-83;
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY	1	CTCCAGGCAATATGTTAACTAATGAAGAAATCTGACCGGACAGCAGTCACTGGGACCTTG	60
Db	366	CTCCAGGCAATATGTTAACTAATGAAGAAATCTGCAACCGCAACGAGTCACTGGGACCTTG	4226
QY	61	CCGTCGATCCTTCCCAACGCTGTACTTTGACGTGAGAGAACTCTGCATATPACTTCAT	120
Db	426	CCGTGCATCCTTCCCAACGCTGTACTTTGACGTGAGAGAACTCTGCATATPACTTCAT	48
QY	121	CTATGAGAGCTGCCGGGSCAATAGAAACAGCTACCGCTCTGAGAGAGGCTGTGATGCTCG	180
Db	486	CTATGAGAGCTGCCGGGSCAATAGAAACAAGTACCGCTCTGAGAGAGGCTGTGATGCTCG	54
QY	181	CTGCTCCGCCACAGAGAAATCCGCCCTGCCCCCTGAGSCAAGGTGGTGTCTTGAC	24
Db	546	CTGCTCCGCCACAGAGAAATCTCCCTGCCCCCTGAGCTCAAGGTGGTGTGTGAC	60
QY	241	GGGAGCTGTGTGTAATGTGTGATCTCTTCTGGAGAGCTCAATGCTTACCTGATTCG	300
Db	606	GGGAGCTGTGTGTAATGTGTGATCTCTTCTGGAGAGCTCAATGCTTACCTGATTCG	66
QY	301	GGTGGCAAGAGAAACCAAGAGCGTGCCTTGCGCAACGTTGAGACTCCGAGAGATGACAA	360
Db	666	GGTGGCAAGAGAAACCAAGAGCGTGCCTTGCGCAACGTTGAGACTCCGAGAGATGACAA	72
QY	361	GGAGCAGCTGCTGTAAGAACACACTATATGCTCGTGA	394
Db	726	GGAGCAGCTGCTGTAAGAACACTATATGCTCGTGA	759

Search completed: January 10, 2003, 02:42:33
Job time : 2223.98 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	862	98.7	1510	19	AAV33063	Human tissue factor
2	843	96.6	1527	24	AA562308	cDNA sequence #95
3	838	96.0	1956	24	AB050508	Human ovarian anti
4	835.4	95.8	4434	24	ABK34480	Human CDNA for nov
5	829	95.0	1330	24	AA570292	Human Placental b
6	829	95.0	1628	20	AA577951	Human ovarian tum
7	825.4	94.5	1578	22	AA57435	Human intestine c
8	825.2	94.5	1544	18	AA570731	Human placental b
9	825.2	94.5	1544	24	ABK84435	Human CDNA differ

PN	WO9833920-A2.
XX	
PD	06-AUG-1998.
XX	
PF	27-JAN-1998; 98WO-US01468

Gen #3430 used to
Human placental co
Human consensus bi
Kuntz type 2 Ser
Human cancer assoc
Human colon cancer
Hepatocyte growth
Human placental bi
Human placental bi
Bladder cancer ass
Human breast cancer
Human breast cancer
Human breast cancer
Gene #0408 used to
Prostate cancer re
Partial human plac
Partial human plac
EST R74593 DNA. H
Human consensus bi
Human ovarian can
cDNA #12 encoding
cDNA #103 encoding
cDNA #332 encoding
Human gene express
Human breast cancer
Human breast cancer
cDNA #1258 encodin
Human gene signatu
Colon adenocarcino
Human ovarian can
Human ovarian carc
Ovarian carcinoma
Partial human plac
Partial human plac
Human ovarian carc
Human ovarian carc
Human ovarian carc

CC invention.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1696 BP; 359 A; 450 C; 510 G; 375 T; 2 other;

Query Match 96.0%; Score 838; DB 24; Length 1696;
Best Local Similarity 98.2%; Pred. No. 2,7e-201;
Matches 858; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

QY 1 CTCACGAGATATGTTCACTATGAAGAAATCTGACCCGCAACGCAAGTCTAGTGGCTTG 60
DB 801 CTCACGAGATATGTTCACTATGAAGAAATCTGACCCGCAACGCAAGTCTAGTGGCTTG 860
QY 61 CCGGCAATCCCTCCACGCTGTTGATCTGAGTGGAGAGAACTCTGCAATTAATCTTCA 120
DB 861 CCGGCAATCCCTCCACGCTGTTGATCTGAGTGGAGAGAACTCTGCAATTAATCTTCA 920
QY 121 CTATGAGAGCTGCGCGGCAATAGAACAGCTACCGCTGAGAGAGCTGATGCTCG 180
DB 921 CTATGAGAGCTGCGCGGCAATAGAACAGCTACCGCTGAGAGAGCTGATGCTCG 980
QY 181 CTGCTCCGCGCAGAGAGAAATCCCTCCCTGCTGCAAAAGTGTGCTTGGC 240
DB 981 CTGCTCCGCGCAGAGAGAAATCCCTCCCTGCTGCAAAAGTGTGCTTGGC 1040
QY 241 GGGGCTGTGCTGATGTTGATCTCTTCTGGAGCTCCATGCTCTGATCTGATCG 300
DB 1041 GGGGCTGTGCTGATGTTGATCTCTTCTGGAGCTCCATGCTCTGATCTGATCG 1100
QY 301 GGTGSCACGAGAGAACACAGACCGCTGCGACCGCTGAGAGCTCCGAGATGACAA 360
DB 1101 GGTGSCACGAGAGAACACAGACCGCTGCGACCGCTGAGAGCTCCGAGATGACAA 1160
QY 361 GAGAGAGCTGTGAAGAAACATATGCTGTACCGCTCTGCGCAAGAGACT GG 419
DB 1161 GAGAGAGCTGTGAAGAAACATATGCTGTACCGCTCTGCGCAAGAGACT GG 1220
QY 420 AAGGAGAGGAGAGACTATGCTGAGCTTTTAAATAGAGAGTTGACTCGGATTTGAGT 479
DB 1221 AAGGAGAGGAGAGACTATGCTGAGCTTTTAAATAGAGAGTTGACTCGGATTTGAGT 1280
QY 480 GATCATTTAGGCTGAGAGCTGTTTCTGCGAGGTTAGAGACGCTCTTCTGCTGCA 539
DB 1281 GATCATTTAGGCTGAGAGCTGTTTCTGCGAGGTTAGAGACGCTCTTCTGCTGCA 1340
QY 540 GGGATGGGTTGCTTGGAAATCCCTGAGAGAGCTCTCTGCGAATGGCTGAGCTG 599
DB 1341 GGGATGGGTTGCTTGGAAATCCCTGAGAGAGCTCTCTGCGAATGGCTGAGCTG 1400
QY 600 CAGCAGCCCGAGAGTTGCTCGCTGATCGATTTCTTCTCGAGTAGAGTTTCTT 659
DB 1401 CAGCAGCCCGAGAGTTGCTCGCTGATCGATTTCTTCTCGAGTAGAGTTTCTT 1460
QY 660 GCTTATGTTGATTCATTTGCTTCTGATCAAGAGTGAATGTTGCTTCT 719
DB 1461 GCTTATGTTGATTCATTTGCTTCTGATCAAGAGTGAATGTTGCTTCT 1520
QY 720 TTTGTTGCTGATTTATGTTTAAATAGAAATTTTAAATTTTATAGATCTG 779
DB 1521 TTTGTTGCTGATTTATGTTTAAATAGAAATTTTAAATTTTATAGATCTG 1580
QY 780 AAGAGAGAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 839
DB 1581 AAGAGAGAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1640
QY 840 TTCAGCATGCTCTTCAAAAAAATTAATTAATTAATTAATTAATTAATTAAT 873
DB 1641 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1674

RESULT 4

ABK34780/c
ID ABK34780 standard; CDNA; 4334 BP.

XX ABK34780;

XX 08-MAY-2002 (first entry)

XX Human CDNA for novel secreted protein, SEQ ID 549.

XX Human; SS; gene; secreted protein; immune deficiency; viral infection;
XX bacterial infection; fungal infection; autoimmune disorder; burn;
XX rheumatoid arthritis; multiple sclerosis; autoimmune thyroiditis;
XX diabetes; allergy; asthma; neurodegenerative disease; Crohn's disease;
XX Alzheimer's disease; Parkinson's disease; liver fibrosis; tumour;
XX coagulation disorder; haemophilia; inflammatory disorder; ulcer;
XX tissue regeneration; wound healing; haematopoiesis; myeloid deficiency;
XX lymphoid cell deficiency.

XX Homo sapiens.

XX W0200177290-A2.

XX 18-OCT-2001.

XX 29-MAR-2001; 2001WO-US10295.

XX 06-APR-2000; 2000US-194941P.

XX (GENY) GENETICS INST INC.

XX Wong GG, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;

XX Gulukota K, Graham JR;

XX WPI; 2002-179323/23.

XX Six hundred and twenty five polynucleotides derived from a variety of
XX human tissue sources which encode secreted proteins, useful for
XX treating immune deficiencies and disorders such as autoimmune disorders

XX Claim 1; Page 293-294; 339pp; English.

XX The invention relates to 625 polynucleotides which have been derived from
XX a variety of human tissue sources and which encode novel secreted
XX proteins, their complements and sequences that hybridise to them.
XX Also included are a vector comprising the polynucleotide, a host cell
XX transformed with the vector, the proteins encoded by the
XX polynucleotides, antibodies that bind to the proteins and identification
XX of modulators of the proteins or the expression of the polynucleotide.
XX The polynucleotides can be used as probes for the identification
XX and isolation of full length cDNA and genomic DNA. The polynucleotides
XX and proteins can also be used as nutritional supplements. The protein
XX is useful in the treatment of various immune deficiencies and disorders
XX such as viral infections, bacterial infections, fungal infections,
XX autoimmune disorders (e.g. rheumatoid arthritis, multiple sclerosis,
XX autoimmune thyroiditis and diabetes) and allergic reactions and
XX conditions (e.g. asthma). They are also useful for treating
XX neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's
XX disease), liver fibrosis, coagulation disorders (e.g. haemophilia),
XX inflammatory disorders (e.g. Crohn's disease) and tumours. They are also
XX useful for tissue regeneration, for wound healing and in the treatment
XX of burns, incisions and ulcers. The proteins are also useful for
XX regulating haematopoiesis, for treating myeloid or lymphoid cell
XX deficiencies. The present sequence is one of the 625 cDNA sequences
XX encoding a secreted protein.

XX Sequence 4334 BP; 1204 A; 1018 C; 891 G; 1221 T; 0 other;

Query Match 95.8%; Score 836.4; DB 24; Length 4334;
Best Local Similarity 98.1%; Pred. No. 1.1e-200;
Matches 857; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

QY 1 CTCACGAGATATGTTCACTATGAAGAAATCTGACCCGCAACGCAAGTCTAGTGGCTTG 60

Db 966 GGTGACGAGGAGAACCCAGAGAGTCCCTGCGACCCGTCGGAGCTCCGAGATGACAA 1025
 Qy 361 GGAGCAGCTGGTGAAGAACCAATATCTCTGACCCGCTCTGCGCAAGAGACT- GGG 419
 Db 1026 GGAGCAGCTGGTGAAGAACCAATATCTCTGACCCGCTCTGCGCAAGAGACTGGGG 1085
 Qy 420 AAGGAGGAGGAGACTATGTGAGCTTTTAAATAGAGGATGACTCGATTTGAGT 479
 Db 1086 AAGGAGGAGGAGACTATGTGAGCTTTTAAATAGAGGATGACTCGATTTGAGT 1145
 Qy 480 GATCATTAAGGCTAGAGTCTGTTCTTGGAGGAGTACGCTCTTGGCTGCA 539
 Db 1146 GATCATTAAGGCTAGAGTCTGTTCTTGGAGGAGTACGCTCTTGGCTGCA 1205
 Qy 540 GGGATGGGTTGCTTTGAAATCCTAGAGGAGCTCCCTCCAGAGCTGCGAGTGG 599
 Db 1206 GGGATGGGTTGCTTTGAAATCCTAGAGGAGCTCCCTCCAGAGCTGCGAGTGG 1265
 Qy 600 CAGCAGCCCGAGTGTCTCTGCTGATGATTTCTTCTCAGAGTGAATTTCTTT 659
 Db 1266 CAGCAGCCCGAGTGTCTCTGCTGATGATTTCTTCTCAGAGTGAATTTCTTT 1325
 Qy 660 GCTTATGTTGAATTCATTCGCTCTTTCTCATCAGAGTGAATTTCTTTCT 719
 Db 1326 GCTTATGTTGAATTCATTCGCTCTTTCTCATCAGAGTGAATTTCTTTCT 1385
 Qy 720 TTGTTGCTGATTTAGTTTAAATGATTAACAAAGTTTAAATGATTTCT 779
 Db 1386 TTGTTGCTGATTTAGTTTAAATGATTAACAAAGTTTAAATGATTTCT 1445
 Qy 780 AAGAGAGGAAAGTAAATGATTAACAAAGTTTAAATGATTTAAAT 839
 Db 1446 AAGAGAGGAAAGTAAATGATTAACAAAGTTTAAATGATTTAAAT 1505
 Qy 840 TTGAGCATGTGCTTCAAAAAA 864
 Db 1506 AAAAAA 1530
 RESULT 6
 AA277551
 ID AA277551 standard; cDNA; 1628 BP.
 AC AA277551;
 DT 10-APR-2000 (first entry)
 DE Human ovarian tumor cDNA library derived EST fragment 102.
 XX
 KM Expressed sequence tag; EST; human; ovarian tumor; anticancer;
 KM gene therapy; treatment; ss.
 XX
 OS Homo sapiens.
 XX
 PN DE19817557-A1.
 XX
 PD 21-OCT-1999.
 XX
 PF 09-APR-1998; 98DE-1017557.
 XX
 PR 09-APR-1998; 98DE-1017557.
 XX
 PA (META-) METABEN GBS GENOMFORSCHUNG MBH.
 XX
 PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
 XX
 DR WPI; 1999-591920/51.
 XX
 PT New nucleic acid sequences expressed in ovarian, and some other, cancer
 PT tissues, and derived polypeptides, for treatment of ovarian cancer and
 PT identification of therapeutic agents
 XX

PS Claim 3; Page 225; 310pp; German.
 XX
 CC This invention describes novel nucleic acid (cDNA) sequences (A) which
 CC have anticancer activity and are highly expressed in ovarian tumor
 CC tissue (and some also in testis and breast cancer tissue). The products
 CC of the invention can be used for gene therapy. (A) are used (i) for
 CC recombinant expression of polypeptides (B) and (ii) to isolate complete
 CC genes. (B) are used (i) to identify agents suitable for treatment of
 CC ovarian cancer; (ii) directly for treating this form of cancer
 CC (including expression from gene therapy vectors) and (iii) for generation
 CC of specific antibodies. (A) are identified by assembling ESTs (expressed
 CC sequence tags) from a particular tissue type before comparison of
 CC expression patterns. This allows a significantly longer fragment of the
 CC gene to be revealed, so should reduce the number of failures associated
 CC with the fact that ESTs from different libraries may represent different
 CC parts of the same unknown gene, distorting the estimated frequency of
 CC occurrence in a particular tissue. AA277450-277572 represent the human
 CC ovarian tumor cDNA library derived EST fragments described in the method
 CC of the invention and encode the protein fragments represented in
 CC AA27505-Y76638.
 XX
 SQ Sequence 1628 BP; 329 A; 433 C; 498 G; 368 T; 0 other;
 Query Match 95.0%; Score 829; DB 20; Length 1628;
 Best Local Similarity 99.2%; Pred. No. 5e-199;
 Matches 849; Conservative 0; Mismatches 15; Indels 1; Gaps 1;
 Qy 1 CTCAGGAGATGTTCACTATGAGAGTACGACCGGCAAGCAGTCACTGAGCTTG 60
 Db 764 CTCAGGAGATGTTCACTATGAGAGTACGACCGGCAAGCAGTCACTGAGCTTG 823
 Qy 61 CCGTCACTCTTCCACGCTGTGATCTTTGACGTGAGAGAGTCTCTGATTAATTCAT 120
 Db 824 CCGTCACTCTTCCACGCTGTGATCTTTGACGTGAGAGAGTCTCTGATTAATTCAT 883
 Qy 121 CTAGAGAGGCTCGCGGGCAATAGAGACGCTTCTGAGAGGCTCTGATGCTCG 180
 Db 884 CTAGAGAGGCTCGCGGGCAATAGAGACGCTTCTGAGAGGCTCTGATGCTCG 943
 Qy 181 CTGCTTCCGCAAGAGAGATCTCCCTGCGCCCTTGGCTCAAGGTGGTGTGCTG 240
 Db 944 CTGCTTCCGCAAGAGAGATCTCCCTGCGCCCTTGGCTCAAGGTGGTGTGCTG 1003
 Qy 241 GGGGCTTCTGATGAGTGTGATCTCTTCTGAGAGCTTCATGCTTACTGATCG 300
 Db 1004 GGGGCTTCTGATGAGTGTGATCTCTTCTGAGAGCTTCATGCTTACTGATCG 1063
 Qy 301 GGTGGACGGAGAGAACAGAGAGCTGCGCGACCGCTGAGAGCTCGGAGATGACA 360
 Db 1064 GGTGGACGGAGAGAACAGAGAGCTGCGCGACCGCTGAGAGCTCGGAGATGACA 1123
 Qy 361 GGAGCAGCTGGTGAAGAACCAATATGCTGTACCGCTCTGCGCAAGAGACT- GGG 419
 Db 1124 GGAGCAGCTGGTGAAGAACCAATATGCTGTACCGCTCTGCGCAAGAGACT- GGG 1183
 Qy 420 AAGGAGGAGGAGACTATGTGAGCTTTTAAATAGAGGATGACTCGATTTGAGT 479
 Db 1184 AAGGAGGAGGAGACTATGTGAGCTTTTAAATAGAGGATGACTCGATTTGAGT 1243
 Qy 480 GATCATTAAGGCTAGAGTCTGTTCTTGGAGGAGTACGCTCTTGGCTGCA 539
 Db 1244 GATCATTAAGGCTAGAGTCTGTTCTTGGAGGAGTACGCTCTTGGCTGCA 1303
 Qy 540 GGGATGGGTTGCTTTGAAATCCTAGAGGAGCTCCCTCCAGAGCTGCGAGTGG 599
 Db 1304 GGGATGGGTTGCTTTGAAATCCTAGAGGAGCTCCCTCCAGAGCTGCGAGTGG 1363
 Qy 600 CAGCAGCCCGAGTGTCTCTGCTGATGATTTCTTCTCAGAGTGAATTTCTTT 659
 Db 1364 CAGCAGCCCGAGTGTCTCTGCTGATGATTTCTTCTCAGAGTGAATTTCTTT 1423
 Qy 660 GCTTATGTTGAATTCATTCGCTCTTTCTCATCAGAGTGAATTTCTTTCT 719

Db 1424 GCTATGTTGAATTCATTGCTCTTTTCATCACAAGATGATGGATGCTTCT 1483
 Qy 720 TTGTTGTCGATTTATGTTTATTAAGATTAACAAAGTTTATTAAGCATTCG 779
 Db 1484 TTGTTGTCGATTTATGTTTATTAAGATTAACAAAGTTTATTAAGCATTCG 1543
 Qy 780 AAAGAGGAAGTAATTAATTAACAGTTTAAATTAAGAGGAGCTTCCCTTTAGATTAAT 839
 Db 1544 AAAGAGGAAGTAATTAATTAACAGTTTAAATTAAGAGGAGCTTCCCTTTAGATTAAT 1603
 Qy 840 TTCAGCATGTGCTTTAAAAAAA 864
 Db 1604 AAAAAAAAAAAAAAAAAAAAAA 1628
 RESULT 7
 AAH57435
 ID AAH57435 standard; cDNA; 1578 BP.
 XX
 AC AAH57435;
 XX
 DT 10-SEP-2001 (first entry)
 XX
 DE Human intestine cell specific cDNA sequence SEQ ID NO:275.
 XX
 KW Human; tissue specific; diagnosis; brain; heart; skeletal muscle;
 KW lung; liver; urens; ovary; stomach; intestine; kidney; pancreas; se;
 KW metabolic disease; developmental disease; cytotoxic; immunomodulatory;
 KW neuropathologic; gene therapy; cancer; immunopathology; neuropathology.
 XX
 OS Homo sapiens.
 XX
 PN WO200132927-A2.
 XX
 PD 10-MAY-2001.
 XX
 PF 02-NOV-2000; 2000MO-US30396.
 XX
 PR 04-NOV-1999; 99US-0163508.
 XX
 PA (INCYTE) INCYTE GENOMICS INC.
 XX
 PI Sornasse T, Seilhamer JJ, Watson GA;
 DR WPI, 2001-291057/30.
 XX
 PT New cell and tissue specific polynucleotides useful for diagnosis,
 PT prognosis or monitoring of treatments for disorders where the gene is
 PT associated with a cancer, immunopathology or neuropathology -
 XX
 PS Claim 1; Page 200-201; 327pp; English.
 XX
 CC AAH57461 to AAH57576 represent cell and tissue specific polynucleotide
 CC sequences (I). (I) can have cytotoxic, immunomodulatory and
 CC neuroprotective activities, and can be used in gene therapy. (I) and
 CC proteins (II) encoded by them are used in high throughput screening
 CC assays to select DNA molecules, RNA molecules, peptide nucleic acids,
 CC mimetics, peptides, proteins, agonists, antagonists, antibodies or
 CC pharmaceutical agents. Expression of (I) in a sample indicates the
 CC differentiation of embryonic stem cells into a tissue selected from
 CC brain, heart, kidney, liver, lung, skeletal muscle or pancreatic
 CC tissues. (I) and (II) are used to produce an expression profile that
 CC defines a metabolic or developmental process, treatment, condition,
 CC disease or disorder. The gene profile can be used for diagnosis,
 CC prognosis or monitoring of treatments and for investigating a
 CC predisposition to a disorder where the gene is associated with a
 CC cancer, immunopathology or neuropathology.
 XX
 SQ Sequence 1578 BP; 303 A; 419 C; 491 G; 365 T; 0 other;
 Query Match 94.5%; Score 825.4; DB 22; Length 1578;
 Best Local Similarity 99.8%; Pred. No. 4e-198;

Matches 837; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 Qy 1 CTCAGGATATGTTCACTATGAAGATATGCTACCGCAACGAGCACTGCGCTTG 60
 Db 737 CTCAGGATATGTTCACTATGAAGATATGCTACCGCAACGAGCACTGCGCTTG 796
 Qy 61 CCGTCACTCTTCCACGCTGCTATCTTTGAAGTGAAGAACTCTCTCAATATCTAT 120
 Db 797 CCGTCACTCTTCCACGCTGCTATCTTTGAAGTGAAGAACTCTCTCAATATCTAT 856
 Qy 121 CTATGAGGCTGCGGGCAATTAAGAAACAGCTACCGCTTGAAGAGGCTGATCTCCG 180
 Db 857 CTATGAGGCTGCGGGCAATTAAGAAACAGCTACCGCTTGAAGAGGCTGATCTCCG 916
 Qy 181 CTGCTTCCGCAAGAGGAATCTCCCTCCCTCCCTCCCTCAAGAGTGTCTCCGCG 240
 Db 917 CTGCTTCCGCAAGAGGAATCTCCCTCCCTCCCTCCCTCAAGAGTGTCTCCGCG 976
 Qy 241 GGGGCTGTTGATGATGATGATCTCTTCTGAGAGGCTCCATGCTCTACCTGATCCG 300
 Db 977 GGGGCTGTTGATGATGATGATCTCTTCTGAGAGGCTCCATGCTCTACCTGATCCG 1036
 Qy 301 GGTGGACGAGAGACAGAGAGCTGCGCCCTGCGACCTGCTGAGCTCCGAGATGCAA 360
 Db 1037 GGTGGACGAGAGACAGAGAGCTGCGCCCTGCGACCTGCTGAGCTCCGAGATGCAA 1096
 Qy 361 GGACACAGCTGTGAATAACACATATGCTGTGACCGCCCTGCGCAAGAGACT- GGG 419
 Db 1097 GGACACAGCTGTGAATAACACATATGCTGTGACCGCCCTGCGCAAGAGACT- GGG 1156
 Qy 420 AAGGAGGAGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 479
 Db 1157 AAGGAGGAGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1216
 Qy 480 GATCATTAGGGCTGAGGTCTGTTCTCTGAGAGTGAAGAGGCTCTTCTGATGCGCA 539
 Db 1217 GATCATTAGGGCTGAGGTCTGTTCTCTGAGAGTGAAGAGGCTCTTCTGATGCGCA 1276
 Qy 540 GGGATGGGTTTGGTATGAAATCCTCTAGAGAGGCTCTCTGAGAGGCTGAGCTGG 599
 Db 1277 GGGATGGGTTTGGTATGAAATCCTCTAGAGAGGCTCTCTGAGAGGCTGAGCTGG 1336
 Qy 600 CAGAGCCCGAGATGTTTCTCTGATGATGATGATGATGATGATGATGATGATGATGAT 659
 Db 1337 CAGAGCCCGAGATGTTTCTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1396
 Qy 660 GCTATGTTGAATTCATTGCTCTTTTCTCATCACAAGATGATGATGATGATGAT 719
 Db 1397 GCTATGTTGAATTCATTGCTCTTTTCTCATCACAAGATGATGATGATGATGAT 1456
 Qy 720 TTGTTGTCGATTTATGTTTATTAAGATTAACAAAGTTTATTAAGCATTCG 779
 Db 1457 TTGTTGTCGATTTATGTTTATTAAGATTAACAAAGTTTATTAAGCATTCG 1516
 Qy 780 AAAGAGGAAGTAATTAATTAACAGTTTAAATTAAGAGGAGCTTCCCTTTAGATTAAT 838
 Db 1517 AAAGAGGAAGTAATTAATTAACAGTTTAAATTAAGAGGAGCTTCCCTTTAGATTAAT 1575
 RESULT 8
 AAT90731
 ID AAT90731 standard; cDNA; 1544 BP.
 XX
 AC AAT90731;
 XX
 DT 20-APR-1998 (first entry)
 XX
 DE Human placental bikunin cDNA.
 XX
 KW Human; placental bikunin; inhibition; trypsin; kallikrein;
 KW plasmin; factor XIIa; treatment; prevention; oedema;
 KW inflammation; infection; granulomatosis; multiple sclerosis;
 KW ischaemia; perioperative blood loss; sepsis; shock; fibrosis;

Dh	846	CTGCTCCGACAGACAGAGAAATCCCTCCCTGACCCCTG3CTCAAAGAGTGGTTCCTGAC	305
Qy	241	GGGGCTGTGTGTGATGATGATTTGATCTCTTCTCTGGAGCTCCATGATCTACTGATCCG	300
Dh	906	GGGGCTGTGTGTGATGATGATTTGATCTCTTCTCTGGAGCTCCATGATCTACTGATCCG	965
Qy	301	GGTGGCACAGAGAAACACAGAGCTGCTCCCTGGACACGCTGTGAGCTCCGAGATGACAA	360
Dh	966	GGTGGCACAGAGAAACACAGAGCTGCTCCCTGGACACGCTGTGAGCTCCGAGATGACAA	1025
Qy	361	GGAGCAGCTGTGTGAAGAACACATATGTCTGTGACCCGCTGTGGCAAGAGACT-966	419
Dh	1026	GGAGCAGCTGTGTGAAGAACACATATGTCTGTGACCCGCTGTGGCAAGAGACT966	1085
Qy	420	AAGGGAGGGAGACTATGTGTGAGCTTTTTTAAATPAGAGATTTGACTCGGATTTGACT	479
Dh	1086	AAGGGAGGGAGACTATGTGTGAGCTTTTTTAAATPAGAGATTTGACTCGGATTTGACT	1145
Qy	480	GATCATTAAGGAGCTGAGTCTCTTCTCTGGAGAGTACGAGCTGCTCTGTCTGTGGCA	539
Dh	1146	GATCATTAAGGAGCTGAGTCTCTTCTCTGGAGAGTACGAGCTGCTCTGTCTGTGGCA	1205
Qy	540	GGGATGGGTTTGGTTTGGAAATCTCTACGAGGCTCTCTGTGATGGCTGTGAGTGG	599
Dh	1206	GGGATGGGTTTGGTTTGGAAATCTCTACGAGGCTCTCTGTGATGGCTGTGAGTGG	1265
Qy	600	CAGCAGCCCCGAGTGTGTTTCTCTGATCATCATTTCTTCTCTCAAGATAGATTTCTTT	659
Dh	1266	CAGCAGCCCCGAGTGTGTTTCTCTGATCATCATTTCTTCTCTCAAGATAGATTTCTTT	1325
Qy	660	GCTTATGTGAATTCATATGCTTC- -TTTCTCATCACAGAGTATGTGGAATCGTTT	717
Dh	1326	GCTTATGTGAATTCATATGCTTCCTCTTCTCTCAACACAGATGATGTGGAATCGTTT	1385
Qy	718	CTTTGTGTGTGATTAATAGTGTGTTTAAATGATTAATAACAAAGTTTATTTATGATTC	777
Dh	1386	CTTTGTGTGTGATTAATAGTGTGTTTAAATGATTAATAACAAAGTTTATTTATGATTC	1445
Qy	778	TGAAGAAGGAAAGTAAATGTACAGTTTAAATAAAGGGGCTTCCCTTTAGATTA	837
Dh	1446	TGAAGAAGGAAAGTAAATGTACAGTTTAAATAAAGGGGCTTCCCTTTAGATTA	1505
Qy	838	ATTTCAGCATGTGCTTTCAAAAAAAAAAAAAA 873	
Dh	1506	ATTTCAGCATGTGCTTTCAAAAAAAAAAAAAA 1542	
RESULT 9			
ABK84435			
ID	ABK84435	standard, cDNA, 1544 BP.	
XX	AC	ABK84435;	
XX	DT	14-AUG-2002 (first entry)	
DE	XX	Human cDNA differentially expressed in granulocytic cells #1006.	
XX	XX	Human, ss; granulocytic cell; DNA chip; bacterial infection;	
KW	XX	viral infection; parasitic infection; protozoal infection;	
KW	XX	fungal infection; sterile inflammatory disease; psoriasis;	
KW	XX	rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;	
KW	XX	cardiac reperfusion injury; renal reperfusion injury; AIDS;	
KW	XX	adult respiratory distress syndrome; inflammatory bowel disease;	
KW	XX	Crohn's disease; ulcerative colitis; periodontal disease;	
KW	XX	granulocyte activation; chronic inflammation; allergy.	
XX	OS	Homo sapiens.	
XX	FN	MO200228999-A2.	
XX	XX	11-APR-2002.	
XX	XX	03-OCT-2001; 2001WO-US30821.	

CC encoded by the present sequence is a Kunitz-type serine protease
 CC inhibitor protein, which can stimulate the rate of mucociliary clearance
 CC of mucus and sputum in lung airways. Therefore, the protein encoded by
 CC the present sequence may be used for treating lung diseases such as CF,
 CC CB, BE, and chronic sinusitis and glue ear which are caused by retention
 CC and accumulation of mucus.

XX
 XX Sequence 1544 BP; 316 A; 393 C; 474 G; 360 T; 1 other;

Query Match 94.3%; Score 823.6; DB 21; Length 1544;
 Best Local Similarity 98.4%; Pred. No. 1.1e-197;

Matches 863; Conservative 0; Mismatches 10; Indels 4; Gaps 3;

QY 1 CTCACGATATGTTCACTATGAAAGAAATCTGCAACCCCAAGCAGCTCATGAGGCTTGG 60
 DB 666 CTCACGATATGTTCACTATGAAAGAAATCTGCAACCCCAAGCAGCTCATGAGGCTTGG 725
 QY 61 CCGTGATCTCTCCCGACCGTGTACTTGTGCGTGAAGAGAACTCTGCAATTAATTCTTAT 120
 DB 726 CCGTGATCTCTCCCGACCGTGTACTTGTGCGTGAAGAGAACTCTGCAATTAATTCTTAT 785
 QY 121 CTATGAGAGCTGCGGGGCAATAGAAACAGCTACCGCTCTGAGAGGCTCATGCTCCG 180
 DB 786 CTATGAGAGCTGCGGGGCAATAGAAACAGCTACCGCTCTGAGAGGCTCATGCTCCG 845
 QY 181 CTGCTCCGCAAGAGAAATCTCTCCGCTTGTGCTCAAGGTGTGTGCTTGGC 240
 DB 846 CTGCTCCGCAAGAGAAATCTCTCCGCTTGTGCTCAAGGTGTGTGCTTGGC 905
 QY 241 GGGGCTGTTCTGATGATGATGATCTCTCTGAGAGCTTCCATGCTCATGCTGATCCG 300
 DB 906 GGGGCTGTTCTGATGATGATGATCTCTCTGAGAGCTTCCATGCTCATGCTGATCCG 965
 QY 301 GGTGGACAGAGAAACAGAGAGCTGCGCCGCAAGCTCTGAGAGCTCCGAGATGACAA 360
 DB 966 GGTGGACAGAGAAACAGAGAGCTGCGCCGCAAGCTCTGAGAGCTCCGAGATGACAA 1025
 QY 361 GAGCAGCTGTGAAAGAACATATGTCTGTGACCGGCTGTGCCAAGAGACT-69G 419
 DB 1026 GAGCAGCTGTGAAAGAACATATGTCTGTGACCGGCTGTGCCAAGAGACTG99G 1085
 QY 420 AAGGAGGGGAGACTATGTGTGAGCTTTTAAATAGAGGATTTGCTCGATTTGACT 479
 DB 1086 AAGGAGGGGAGACTATGTGTGAGCTTTTAAATAGAGGATTTGCTCGATTTGACT 1145
 QY 480 GATCATTAGGGGCTGAGCTGTTCTCTGAGAGGATGAGAGGCTCTCTGCTGCTGCA 539
 DB 1146 GATCATTAGGGGCTGAGCTGTTCTCTGAGAGGATGAGAGGCTCTCTGCTGCTGCA 1205
 QY 540 GGGATGGGTTGCTTTGAAATCCTCTAGAGAGCTCTCTCGCATGGCTGAGTCTGG 599
 DB 1206 GGGATGGGTTGCTTTGAAATCCTCTAGAGAGCTCTCTCGCATGGCTGAGTCTGG 1265
 QY 600 CAGCAGCCCGAGTGTGTTCTCTGCTGATGCAATTTCTTCTCCAGGTAGAGTTTCTTT 659
 DB 1266 CAGCAGCCCGAGTGTGTTCTCTGCTGATGCAATTTCTTCTCCAGGTAGAGTTTCTTT 1325
 QY 660 GCTTATGTGAATTCATGCT 717
 DB 1326 GCTTATGTGAATTCATGCT 1385
 QY 718 CTTTGTGTTGCTGATTTATGTTTAAATAGAAAGAGTTTATTAAGATTC 777
 DB 1386 CTTTGTGTTGCTGATTTATGTTTAAATAGAAAGAGTTTATTAAGATTC 1445
 QY 778 TGAAGAGAGAAAGTAAATGTAAGATTAAATTAATTAATTAATTAATTAATTAATTAAT 837
 DB 1446 TGAAGAGAGAAAGTAAATGTAAGATTAAATTAATTAATTAATTAATTAATTAATTAAT 1505
 QY 838 ATTT-CAGCATGTGCTTCAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAAT 873
 DB 1506 ATTTCAAGCATGTGCTTCAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1542

RESULT 12

AAT90733
 ID AAT90733 standard; CDNA; 1528 BP.

XX
 AC AAT90733;

XX
 DT 20-APR-1998 (first entry)

XX
 DE Human consensus bikunin cDNA.

XX
 KM Human; consensus bikunin; inhibition; trypsin; kallikrein;

XX
 KM plasmin; factor Xlla; treatment; prevention; oedema;

XX
 KM inflammation; infection; granulomatosis; multiple sclerosis;

XX
 KM ischaemia; perioperative blood loss; sepsis; shock; fibrosis;

XX
 KM blood coagulation disease; polytrauma; stroke; haemorrhage;

XX
 KM gastric cancer; cervical cancer; metastasis; blood loss; ss.

XX
 OS Homo sapiens.

XX
 FH Key Location/Qualifiers

XX
 FH CDS 313..1059

XX
 FT /tag= a

XX
 FT /product= bikunin

XX
 FT unsure 313..315

XX
 FT /tag= b

XX
 FT /note= "encodes Met"

XX
 PA (PARB) BAYER CORP.

XX
 PI Davis G, Delaria KA, Marlor CW, Muller DK, Tamburini PP;

XX
 DR WPI; 1997-470876/43.

XX
 DR P-PSDB; AAW30044.

XX
 PS Claim 4; Fig 4C; 110pp; English.

XX
 CC The present sequence encodes a consensus human bikunin, which

XX
 CC inhibits, e.g. trypsin, Kallikrein, plasmin and factor Xlla.

XX
 CC Bikunin can be used to treat or prevent brain and spinal cord

XX
 CC oedema, inflammation, infection or granulomatosis, multiple

XX
 CC sclerosis, ischaemia, perioperative blood loss, sepsis, shock,

XX
 CC fibrosis, blood coagulation diseases, polytrauma, stroke,

XX
 CC cerebral or subarachnoid haemorrhage and gastric or cervical

XX
 CC cancer and prevent metastasis. It is particularly useful for

XX
 CC reducing blood loss during surgery, and can also be used to treat

XX
 CC other cancer, arthritis, anaemia, non-insulin dependent diabetes,

XX
 CC influenza and similar viral infections, acute pancreatitis and

XX
 CC gout, and prevent pre-term labour. It has similar properties to

XX
 CC immunogenic, but is less highly charged so should be less

XX
 CC altered. It also reduces or eliminates the need for whole donor

XX
 CC blood or blood products during surgery, thereby reducing the risk

XX
 CC of infection and other adverse side effects, as well as reducing

XX
 CC the cost of surgery.

XX
 SQ Sequence 1528 BP; 320 A; 386 C; 467 G; 352 T; 3 other;

Query Match 92.6%; Score 808; DB 18; Length 1528;
 Best Local Similarity 97.5%; Pred. No. 9,7e-194;
 Matches 842; Conservative 0; Mismatches 20; Indels 2; Gaps 2;

QY 1 CTCGAGCATATGTTCACTATGAAGTAATACGACCCGCAAGCAGCTACCTGGGCTTG 60
 DB 666 CTCGAGCATATGTTCACTATGAAGTAATACGACCCGCAAGCAGCTACCTGGGCTTG 725
 QY 61 CCGTGACATCTTCCACGCTGTGATCTTGAAGTGAAGGAACTCTCGAATACTTCAT 120
 DB 726 CCGTGACATCTTCCACGCTGTGATCTTGAAGTGAAGGAACTCTCGAATACTTCAT 785
 QY 121 CTATGAGGCTCCGCGGCAATTAAGAACGCTACCGCTTGAAGAGGCTCGATCTCCG 180
 DB 786 CTATGAGGCTCCGCGGCAATTAAGAACGCTACCGCTTGAAGAGGCTCGATCTCCG 845
 QY 181 CTGCTCCGCGAGAGAGAAATCTCCCTGCTGGCTGCAAAAGTGTGCTTGGC 240
 DB 846 CTGCTCCGCGAGAGAGAAATCTCCCTGCTGGCTGCAAAAGTGTGCTTGGC 905
 QY 241 GGGGCTGTTCGTATGATGTTGATCTCTCTGGAGCTCCATGATCTGATCCG 300
 DB 906 GGGGCTGTTCGTATGATGTTGATCTCTCTGGAGCTCCATGATCTGATCCG 965
 QY 301 GGTGCAAGGAGAAACAGAGAGCTGCTCCGCACTGCTGAGAGCTCCGAGATGCAA 360
 DB 966 GGTGCAAGGAGAAACAGAGAGCTGCTCCGCACTGCTGAGAGCTCCGAGATGCAA 1025
 QY 361 GGAGCAGCTGTGTAAGAACATATGCTGTAAGCTGCTGCGCAAGAGAGCT-866 419
 DB 1026 GGAGCAGCTGTGTAAGAACATATGCTGTAAGCTGCTGCGCAAGAGAGCT-866 1085
 QY 420 AAGGAGAGGAGAACTATGTGTGAAGCTTTTAAATGAAGGATTTGATGAT 479
 DB 1086 AAGGAGAGGAGAACTATGTGTGAAGCTTTTAAATGAAGGATTTGATGAT 1145
 QY 480 GATCATTAGGCTGAGCTGCTTTCTCTGGAGGTAGAGAGCTGCTTCTGGTGGCA 539
 DB 1146 GATCATTAGGCTGAGCTGCTTTCTCTGGAGGTAGAGAGCTGCTTCTGGTGGCA 1205
 QY 540 GGGATGAGCTTTGTTGAAATCTCTAGAGAGCTCTCTGCAAGCTGAGTGG 599
 DB 1206 GGGATGAGCTTTGTTGAAATCTCTAGAGAGCTCTCTGCAAGCTGAGTGG 1265
 QY 600 CAGCAGCCCGAGTGTCTCTGCTGCTGATGCAATTTCTTCTCAGAGTGAATTTCTT 659
 DB 1266 CAGCAGCCCGAGTGTCTCTGCTGCTGATGCAATTTCTTCTCAGAGTGAATTTCTT 1325
 QY 660 GCTTATGTTGAATTCATTCGCTTCTCTGATCAGAGAGTATGTTGATGCTTCT 719
 DB 1326 GCTTATGTTGAATTCATTCGCTTCTCTGATCAGAGAGTATGTTGATGCTTCT 1384
 QY 720 TTTGTTGCTGATTTAGTTTCTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 779
 DB 1385 TTTGTTGCTGATTTAGTTTCTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1444
 QY 780 AAGAGAGAGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 839
 DB 1445 AAGAGAGAGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1504
 QY 840 TTCAGCATGCTTCTTCAAAAAA 863
 DB 1505 AAAAAA 1528

RESULT 13
 ID ABA83112
 AC ABA83112;
 XX 08-FEB-2002 (first entry)
 DT XX

DE Kunitz type 2 Ser protease inhibitor ovarian tumour marker gene, #61.
 XX
 KW Ovarian tumour marker gene; human; overexpression; upregulation;
 KW epithelial tumour; cancer; diagnosis; prognosis; disease monitoring;
 KW identification; serous cystadenoma; borderline serous tumour;
 KW mucinous cystadenocarcinoma; mucinous mucinous tumour; endometrioid carcinoma;
 KW undifferentiated carcinoma; clear cell adenocarcinoma; cystadenofibroma;
 KW adenofibroma; Brenner tumour; serial analysis of gene expression; SAGE;
 KW immune response pathway; cell proliferation regulation; protein folding;
 KW membrane localised; secreted; therapeutic target; cytostatic;
 KW gene therapy; vaccine; ds.
 OS
 XX Homo sapiens.
 PN W02001.75177-A2.
 PD 11-OCT-2001.
 XX
 XX 03-APR-2001; 2001WO-US10947.
 PF 03-APR-2000; 2000US-194336P.
 PR (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA Morin PJ, Sherman-Baust CA, Pizer ES, Hough CD;
 PI WPI; 2001-626450/72.
 DR P-PSDB; ABB50286.
 XX
 XX Detecting and identifying ovarian tumor, identifying increased risk for
 PT developing ovarian cancer, and determining effectiveness of ovarian
 PT cancer treatment, by measuring expression level of ovarian tumor marker
 PT gene -
 PS Claim 23; Page 104-105; 140pp; English.
 XX
 XX The invention relates to methods for diagnosing and prognosing ovarian
 CC tumors in an individual via the detection and measurement of the
 CC expression of ovarian tumor marker genes (ABA83081-ABA83122, ABA83180,
 CC ABA83182 and ABA83184) or segments thereof (ABA83123-ABA83169, ABA83179,
 CC ABA83181 and ABA83183). The methods of the invention are useful for
 CC detecting an ovarian tumour in a patient, for identifying an individual
 CC at increased risk for developing ovarian cancer, in prognostic tests for
 CC assessing the relative severity of ovarian cancer, in tests for
 CC monitoring a patient in remission from ovarian cancer, in tests for
 CC monitoring disease status in a patient being treated for ovarian cancer.
 CC The methods can additionally be used to identify a particular tumour as
 CC being an ovarian tumour (i.e., an epithelial ovarian tumour selected from
 CC serous cystadenoma, borderline serous tumour, serous cystadenocarcinoma,
 CC mucinous cystadenoma, borderline mucinous tumour, mucinous
 CC cystadenocarcinoma, endometrioid carcinoma, undifferentiated carcinoma,
 CC clear cell adenocarcinoma, cystadenofibroma, adenofibroma and Brenner
 CC tumour). The ovarian tumour marker genes of the invention were identified
 CC using SAGE (serial analysis of gene expression) and were found to be
 CC overexpressed in a broad variety of ovarian epithelial tumour cells
 CC relative to normal ovarian epithelial cells. The marker genes are
 CC implicated in immune response pathways, in the regulation of cell
 CC proliferation and in protein folding, and many of these are membrane-
 CC localised or secreted. In addition to their use as diagnostic and
 CC prognostic markers, the ovarian tumour marker genes or their encoded
 CC proteins may be used as therapeutic targets for the treatment and
 CC prevention of ovarian cancer. Sequences ABA83081-ABA83122, ABA83180,
 CC ABA83182 and ABA83184 represent the ovarian tumour marker genes of
 CC the invention.
 CC
 XX
 XX Sequence 1564 BP; 297 A; 419 C; 491 G; 357 T; 0 other;
 SQ

Query Match 92.3%; Score 805.6; DB 22; Length 1564;
 Best Local Similarity 99.3%; Pred. No. 3.9e-193;
 Matches 830; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY 1 CTCGAGCATATGTTCACTATGAAGTAATACGACCCGCAAGCAGCTACCTGGGCTTG 60


```

OY 301 GATGACAGGAGGAAACAGAGACCGTGCCTGCGACCGTCTGAGCTCCGAGATGACAA 360
DB 1084 GGTGGACAGGAGGAAACAGAGACCGTGCCTGCGACCGTCTGAGCTCCGAGATGACAA 1143
OY 361 GAGACAGCTGCTGAAACACATATGTCCTGTGACCGCCCTGTGCGCAAGAGATGGGA 420
DB 1144 GAGACAGCTGCTGAAACACATATGTCCTGTGACCGCCCTGTGCGCAAGAGATGGGA 1202
OY 421 AGGAGGGGAGACTATGCTGAGCTTTTAAATAGAGGATTCACCGATTTGAGTG 480
DB 1203 AAGGAGGGGAGACTATGCTGAGCTTTTAAATAGAGGATTCACCGATTTGAGTG 1261
OY 481 ATCATTAAGGCTGAGCTGTTCTCTGAGAGGATGAGAGGCTGCTCTGCTGCTGAG 540
DB 1262 ATCATTAAGGCTGAGCTGTTCTCTGAGAGGATGAGAGGCTGCTCTGCTGCTGAG 1321
OY 541 GATGAGGCTTTGCTTTGAAATCTCTGAGAGGCTCTCTGCTGCTGAGCTGAG 600
DB 1322 GATGAGGCTTTGCTTTGAAATCTCTGAGAGGCTCTCTGCTGCTGAGCTGAG 1381
OY 601 AGCAGCCCGAGTGTCTCTGAGATGATTTCTTCTCTGAGTGAATTTCTTTG 660
DB 1382 AGCAGCCCGAGTGTCTCTGAGATGATTTCTTCTCTGAGTGAATTTCTTTG 1440
OY 661 CTATGTTGAATTCATTCCTCTTTCTCATACAGAAATGATGTTGAAATGCTTT 720
DB 1441 CTATGTTGAATTCATTCCTCTTTCTCATACAGAAATGATGTTGAAATGCTTT 1499
OY 721 TTGTTTGTGATTTATGTTTATTAAGTATTAACAAAGTTTATTAAGCTTTCTGA 780
DB 1500 TTGTTTGTGATTTATGTTTATTAAGTATTAACAAAGTTTATTAAGCTTTCTGA 1559
OY 781 AAGAGAGAAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 840
DB 1560 AAGAGAGAAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1618
OY 841 TCAGCATGTGCTTCAAAAAAAAAAAAAA 871
DB 1619 AAAAAAAAAAAAAAAAAAAAAA 1649

RESULT 15
AA16487
ID AA16487 standard; DNA; 597 BP.
XX
AC AA16487;
XX
DT 14-JUN-2000 (first entry)
XX
DE Human colon cancer differentially expressed nucleotide sequence #492.
XX
KM Colon cancer; detect; differential expression; human; treatment;
XX
KW detect mutation; non-invasive diagnostic method; ds.
XX
OS Homo sapiens.
XX
PN WO200012702-A2.
XX
PD 09-MAR-2000.
XX
PF 30-AUG-1999; 99WO-US19424.
XX
PR 31-AUG-1998; 98US-0098639.
XX
PR 27-JAN-1999; 99US-0113993.
XX
PA (FARB) BAYER CORP.
XX
PI Endege NO, Steimann KE, Astle JH, Burgess CC, Carroll E,
PI Catino T, Dwivedi P, Ford DW, Lewis ME, Molino GA, Monahan JB,
PI Schlegel R;
XX
DR WPI, 2000-256641/22.

```

```

XX
PT Novel nucleic acids and proteins for identifying therapeutic agents
PT useful for treating and diagnosing cancer, especially colon cancer -
XX
PS Claim 16; Page 323; 345pp; English.
XX
CC This sequence represents a human nucleotide sequence which is
CC differentially expressed in colon cancer cells compared to the expression
CC levels in normal cells. The nucleotide sequence can be used as a source
CC of primers and probes. The nucleotide sequence is useful for determining
CC the phenotype of a cell by detecting the differential expression of the
CC sequence relative to a normal cell. The probes derived from the sequence
CC can also be used to determine the phenotype of cells in a sample. Probes
CC and antibodies which hybridize to the nucleotide sequence can also be
CC used to determine the phenotype of a cell. The primers are useful for
CC detecting a mutation in a test nucleotide sequence and also for detecting
CC cancer, preferably colon cancer. Antibodies against the protein encoded
CC by the nucleotide sequence can also be used in a method to detect colon
CC cancer. The diagnostic method is non-invasive and accurate for diagnosing
CC colon cancer at an early stage.
XX
SQ Sequence 597 BP; 121 A; 117 C; 157 G; 162 T; 40 other;
XX
Query Match 48.7%; Score 425.4; DB 21; Length 597;
Best Local Similarity 93.3%; Pred. No. 2.e-97;
Matches 460; Conservative 0; Mismatches 31; Indels 2; Gaps 2;
OY 300 GGTGGACAGGAGGAAACAGAGACCGTGCCTGCGACCGTCTGAGCTCCGAGATGAC 359
DB 8 GGTGGACAGGAGGAAACAGAGACCGTGCCTGCGACCGTCTGAGCTCCGAGATGAC 67
OY 360 AGGAGAGCTGCTGAAAGACATATGTCCTGTGACCGCCCTGTGCGCAAGAGACT-GG 418
DB 68 AGGAGAGCTGCTGAAAGACATATGTCCTGTGACCGCCCTGTGCGCAAGAGACTGG 127
OY 419 GAAGGAGAGGAGACTATGCTGAGCTTTTAAATAGAGGATTTGACTCGGATTTAG 478
DB 128 GAAGGAGAGGAGACTATGCTGAGCTTTTAAATAGAGGATTTGACTCGGATTTAG 187
OY 479 TGATCATTAAGGCTGAGCTGTTTCTCTGAGAGGTAGAGCGCTGCTGCTGCTGGC 538
DB 188 TGATCATTAAGGCTGAGCTGTTTCTCTGAGAGGTAGAGCGCTGCTGCTGCTGGC 247
OY 539 AGGAGAGGCTTTGCTTTGAAATCTCTAGAGGCTCTCTGAGCTGCTGAGCTG 598
DB 248 ANGAGAGGCTTTGCTTTGAAATCTCTAGAGGCTCTCTGAGCTGCTGAGCTG 307
OY 599 GCAGCAGCCCGAGTGTCTCTGAGAGGCTCTCTGAGCTGCTGAGCTGCTG 658
DB 308 GCAGCAGCCCGAGTGTCTCTGAGAGGCTCTCTGAGCTGCTGAGCTGCTG 367
OY 659 TGCTTATGTTGAATTCATTCCTTTCTCATACAGAAATGATGTTTGAATCGTTT 717
DB 368 TGCTTATGTTGAATTCATTCCTTTCTCATACAGAAATGATGTTTGAATCGTTT 427
OY 718 CTTTGTGTTGCTGATTTATGCTTTTAAATAGATTAACAAAGTTTATTAAGCTTC 777
DB 428 CTTTGTGTTGCTGATTTATGCTTTTAAATAGATTAACAAAGTTTATTAAGCTTC 487
OY 778 TGAAGAGAGGAA 790
DB 488 TTAANAAGGAA 500

```

Search completed: January 10, 2003, 01:53:34
Job time : 261.933 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 10, 2003, 01:47:01 (Search time 51.01 Seconds

(without alignments)
5248.554 Million cell updates/sec

Title: US-09-825-682A-56

Perfect score: 873

Sequence: 1 ctccgcgcatatgttcaact.....tcaaaaaaaaaaaaaaaaaa 873

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/ina/5A.COMB.seq: *
4: /cgn2_6/ptodata/1/ina/5B.COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PTCDS.COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	862	98.7	1610	4	US-09-013-896A-1
2	425.4	48.7	1597	4	US-09-385-982-492
3	394	45.1	759	1	US-08-685-660A-4
4	394	45.1	759	2	US-08-974-196-4
5	282.4	32.3	287	4	US-09-013-896A-13
6	260	29.8	273	4	US-09-013-896A-14
7	250	28.6	256	4	US-09-013-896A-15
8	188	21.5	201	4	US-09-013-896A-17
9	167.4	19.2	276	4	US-09-404-879A-169
10	167	19.1	207	4	US-09-404-879A-356
11	167	19.1	371	4	US-09-404-879A-365
12	165.4	18.9	276	4	US-09-404-879A-168
13	146.8	16.8	199	4	US-09-222-575-125
14	126.8	8.3	783	4	US-09-020-956-15
15	126.8	8.3	783	4	US-09-030-607-15
16	126.8	8.3	783	4	US-09-605-785-15
17	126.8	8.3	783	4	US-09-439-313-15
18	126.8	8.3	783	4	US-09-352-616A-15
19	126.8	8.3	783	4	US-09-232-149A-15
20	126.8	8.3	783	4	US-09-232-149A-15
21	71.4	8.2	1542	4	US-08-685-558A-8
22	71.4	8.2	1542	4	US-08-685-558A-8
23	68.6	7.9	1870	4	US-09-765-449-8
24	56.8	6.5	192	6	US-09-071-709-6
25	56.8	6.5	192	6	US-09-071-709-6
26	56.8	6.5	192	6	US-09-071-709-6
27	56.6	6.5	245	6	US-09-071-709-6

28	56.6	6.5	245	6	US-09-071-709-6	Patent No. 5223482
29	55.8	6.4	704	2	US-08-829-876-100	Sequence 100, App
30	55.8	6.4	704	2	US-08-829-876-104	Sequence 104, App
31	55.8	6.4	704	2	US-09-234-874A-100	Sequence 100, App
32	55.8	6.4	704	2	US-09-234-874A-104	Sequence 104, App
33	55	6.3	185	4	US-08-829-876-76	Sequence 76, App
34	55	6.3	185	4	US-09-234-874A-76	Sequence 76, App
35	55	6.3	197	2	US-08-829-876-78	Sequence 78, App
36	55	6.3	197	2	US-09-234-874A-78	Sequence 78, App
37	55	6.3	245	6	US-09-071-709-6	Patent No. 5187153
38	55	6.3	445	2	US-08-829-876-80	Sequence 80, App
39	55	6.3	445	2	US-08-829-876-82	Sequence 82, App
40	55	6.3	445	2	US-08-829-876-84	Sequence 84, App
41	55	6.3	445	2	US-08-829-876-86	Sequence 86, App
42	55	6.3	445	2	US-08-829-876-88	Sequence 88, App
43	55	6.3	445	2	US-08-829-876-96	Sequence 96, App
44	55	6.3	445	2	US-08-829-876-98	Sequence 98, App
45	55	6.3	445	2	US-08-829-876-106	Sequence 106, App

ALIGNMENTS

RESULT 1
US-09-013-896A-1
Sequence 1, Application US/09013896A
Patent No. 6262233
GENERAL INFORMATION:
APPLICANT: GENTZ, REINER
TITLE OF INVENTION: TISSUE FACTOR PATHWAY INHIBITOR-3
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSER: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVE., NW, STB. 600
CITY: WASHINGTON
STATE: DC
COUNTRY: US
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/013,896A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.1290001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1610 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 361..1116
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 361..439
NAME/KEY: mat_peptide
LOCATION: 442..1116
US-09-013-896A-1
Query Match 98.7%, Score 862, DB 4, Length 1610;

Best Local Similarity 99.9%; Pred. No. 1e-211; Matches 873; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```

QY 1 CTCACGATATGTTCACTATGAAATCTGCAACCGCAACGAGCTCGAGGCTTG 60
Db 726 CTCACGATATGTTCACTATGAAATCTGCAACCGCAACGAGCTCGAGGCTTG 785
QY 61 CCGTGATCCTTCCACCGCTGTACTTTGACGTGAGAGAACTCTGCAATTAATTCT 120
Db 786 CCGTGATCCTTCCACCGCTGTACTTTGACGTGAGAGAACTCTGCAATTAATTCT 845
QY 121 CTATGAGAGCTCCCGGCAATTAAGAACACTACCGCTCTGAGAGGCTGACCTCG 180
Db 846 CTATGAGAGCTCCCGGCAATTAAGAACACTACCGCTCTGAGAGGCTGACCTCG 905
QY 181 CTGCTTCCGCAAGAGAAATCTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Db 906 CTGCTTCCGCAAGAGAAATCTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 965
QY 241 GGGGCTGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
Db 966 GGGGCTGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1025
QY 301 GGTGGACGAGAGAACAGAGAGCTGCTGCGACCTGCTGAGCTCGAGATGACAA 360
Db 1026 GGTGGACGAGAGAACAGAGAGCTGCTGCGACCTGCTGAGCTCGAGATGACAA 1085
QY 361 GAGACACTGCTGAAGAACATATGTCCTGTGACCGCCCTTCCGCAAGAGCT- 419
Db 1086 GAGACACTGCTGAAGAACATATGTCCTGTGACCGCCCTTCCGCAAGAGCT- 1145
QY 420 AAGGAGAGGAGACTATGAGCTTTTAAATAGAGGATGATCGGATTTGAGT 479
Db 1146 AAGGAGAGGAGACTATGAGCTTTTAAATAGAGGATGATCGGATTTGAGT 1205
QY 480 GATCATTAGAGGCTGAGGCTGCTTCTGAGAGGTAGAGAGGCTCTTCTGATG 539
Db 1206 GATCATTAGAGGCTGAGGCTGCTTCTGAGAGGTAGAGAGGCTCTTCTGATG 1265
QY 540 GGGATGGGTTTGTGAAATCTCTGAGAGGTCTCTCTGCAATGAGCTGAGTCTG 599
Db 1266 GGGATGGGTTTGTGAAATCTCTGAGAGGTCTCTCTGCAATGAGCTGAGTCTG 1325
QY 600 CAGAGCCCGAGATGTTCTCTGAGATGATTTCTTCCAGAGTAGATTTCTT 659
Db 1326 CAGAGCCCGAGATGTTCTCTGAGATGATTTCTTCCAGAGTAGATTTCTT 1385
QY 660 GCTTATGTTGAATTCATTCCTCTTCTCATCAGAGATGATGGAATGCTTCT 719
Db 1386 GCTTATGTTGAATTCATTCCTCTTCTCATCAGAGATGATGGAATGCTTCT 1445
QY 720 TTTGTTGCTGATTTATGTTTAAAGTAAACAAAGTTTAAATAGATCTG 779
Db 1446 TTTGTTGCTGATTTATGTTTAAAGTAAACAAAGTTTAAATAGATCTG 1505
QY 780 AAAGAGAGAAAGTAAAGTAAAGTAAAGAGGAGGCTTCCCTTAGAATAAT 839
Db 1506 AAAGAGAGAAAGTAAAGTAAAGTAAAGAGGAGGCTTCCCTTAGAATAAT 1565
QY 840 TTCAGCATGTGCTTTCAAAAA 873
Db 1566 TTCAGCATGTGCTTTCAAAAA 1599

```

RESULT 2
US-09-385-982-492
; Sequence 492, Application US/09385982
; Patent No. 6262314
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCNNA-260XX

CURRENT APPLICATION NUMBER: US/09/385,982
; CURRENT FILING DATE: 1999-08-30
; EARLIER APPLICATION NUMBER: 09/328,111
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: 60/117,393
; EARLIER FILING DATE: 1998-01-27
; EARLIER APPLICATION NUMBER: 60/098,639
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 492
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc.feature
; LOCATION: (1)..(597)
; OTHER INFORMATION: n = A,T,C or G
US-09-385-982-492

Query Match 48.7%; Score 425.4; DB 4; Length 597;
Best Local Similarity 93.3%; Pred. No. 5.6e-100;
Matches 460; Conservative 0; Mismatches 31; Indels 2; Gaps 2;

```

QY 300 GGGTGGACGAGAGAACAGAGAGCTGCTGCGACCTGCTGAGCTCGAGATGACA 359
Db 8 GGGTGGACGAGAGAACAGAGAGCTGCTGCGACCTGCTGAGCTCGAGATGACA 67
QY 360 AGAGAGAGCTGAGAGAACACATATGCTGTGACCGCCCTGCGCAAGAGACT- 418
Db 68 AGAGAGAGCTGAGAGAACACATATGCTGTGACCGCCCTGCGCAAGAGACT- 127
QY 419 GAAAGAGAGGAGACTATGAGCTTTTAAATAGAGGATGATCGATTTGAG 478
Db 128 GAAAGAGAGGAGACTATGAGCTTTTAAATAGAGGATGATCGATTTGAG 187
QY 479 TGATCATTAGAGGCTGAGGCTGCTTCTGAGAGGTAGAGAGGCTCTTCTGATG 538
Db 188 TGATCATTAGAGGCTGAGGCTGCTTCTGAGAGGTAGAGAGGCTCTTCTGATG 247
QY 539 AGGATGGGTTTCTTGGAAATCTCTAGAGAGCTCTCTCGATGAGCTGAGCTG 598
Db 248 AGGATGGGTTTCTTGGAAATCTCTAGAGAGCTCTCTCGATGAGCTGAGCTG 307
QY 599 GAGCAGACCCCGAGTGTCTCTGAGATGATTTCTTCCAGAGTAGATTTCTT 658
Db 308 GAGCAGACCCCGAGTGTCTCTGAGATGATTTCTTCCAGAGTAGATTTCTT 367
QY 659 TGCTTATGTTGAATTCATTCCTTCTCATCAGAGATGATGAT- GGAATGCTT 717
Db 368 TGCTTATGTTGAATTCATTCCTTCTCATCAGAGATGATGATGATGGAATGCTT 427
QY 718 CTTTGTGTTGCTGATTTATGTTTAAAGTAAACAAAGTTTAAATAGATCTG 777
Db 428 CTTTGTGTTGCTGATTTATGTTTAAAGTAAACAAAGTTTAAATAGATCTG 487
QY 778 TGAAGAAGGAAA 790
Db 488 TGAAGAAGGAAA 500

```

RESULT 3
US-08-685-660A-4
; Sequence 4, Application US/0868560A
; Patent No. 5731412
; GENERAL INFORMATION:
; APPLICANT: SHIMOWURA, Takeshi
; APPLICANT: KAWAGUCHI, Toshiya
; APPLICANT: KITAMURA, Naomi
; TITLE OF INVENTION: NOVEL PROTEIN, DNA CODING FOR SAME
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:

```

ADDRESSER: SUGHRUE, MION, ZINN, MACPEAK & SEAS
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20037

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/685,660A
FILING DATE: 24-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JPA Hei 7-187134
FILING DATE: 24-JUL-1995
ATTORNEY/AGENT INFORMATION:
NAME: KIT, Gordon
REGISTRATION NUMBER: 30,764
REFERENCE/DOCKET NUMBER: Q-42295
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 759 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: linear
ANTISENSE: no
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
STRAIN: MKN45
FEATURE:
NAME/KEY: coding sequence
LOCATION: 1 to 759
IDENTIFICATION METHOD: by experiment
NAME/KEY: signal peptide
LOCATION: 1 to 81
IDENTIFICATION METHOD: by experiment
NAME/KEY: mature peptide
LOCATION: 82 to 759
IDENTIFICATION METHOD: by experiment
US-08-685-660A-4

Query Match 45.1%; Score 394; DB 1; Length 759;
Best Local Similarity 100.0%; Pred. No. 7.1e-92;
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCAGGATATGTTCACTATGAAGATACTGACCGCAACGAGTCACTGAGGCTTG 60
DB 366 CTCAGGATATGTTCACTATGAAGATACTGACCGCAACGAGTCACTGAGGCTTG 425
QY 61 CCGTGCACTCTTCCACGCTGTACTTTGAGAGTGAGAGAACTCTTGCAATACTTCA 120
DB 426 CCGTGCACTCTTCCACGCTGTACTTTGAGAGTGAGAGAACTCTTGCAATACTTCA 485
QY 121 CTATGAGAGCTGCGGAGGCAATTAAGACAGCTACCGCTCTGAGAGAGGCTCTGATGCTCG 180
DB 486 CTATGAGAGCTGCGGAGGCAATTAAGACAGCTACCGCTCTGAGAGAGGCTCTGATGCTCG 545
QY 181 CTGCTTCCGACAGAGGAATCTCCCTGCTGCTTGGCTCAAGAGTGAGTGTCTTGAG 240
DB 546 CTGCTTCCGACAGAGGAATCTCCCTGCTGCTTGGCTCAAGAGTGAGTGTCTTGAG 605
QY 241 GGGGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
DB 606 GGGGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 665
QY 301 GGTGGACGAGGAAACAGAGAGCTGCGCGCAGCCGTCTGAGACTCGGAGATGACAA 360

```

```

DB 666 GGTGGACGAGGAAACAGAGAGCTGCGCGCAGCCGTCTGAGACTCGGAGATGACAA 725
QY 361 GGAGCAGCTGTGTGAAGAACATATGCTCTGTGA 394
DB 726 GGAGCAGCTGTGTGAAGAACATATGCTCTGTGA 759

RESULT 4
US-08-974-196-4
Sequence 4, Application US/08974196
Patent No. 5854396
GENERAL INFORMATION:
APPLICANT: SHIMOMURA, Takeshi
APPLICANT: KAWAGUCHI, Toshiya
APPLICANT: KITAMURA, Naomi
TITLE OF INVENTION: NOVEL PROTEIN, DNA CODING FOR SAME
TITLE OF INVENTION: AND METHOD OF PRODUCING THE PROTEIN
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSER: SUGHRUE, MION, ZINN, MACPEAK & SEAS
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,196
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/685,660
FILING DATE: 24-JUL-1996
APPLICATION NUMBER: JPA Hei 7-187134
FILING DATE: 24-JUL-1995
ATTORNEY/AGENT INFORMATION:
NAME: KIT, Gordon
REGISTRATION NUMBER: 30,764
REFERENCE/DOCKET NUMBER: Q-42295
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 759 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: linear
ANTISENSE: no
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
STRAIN: MKN45
FEATURE:
NAME/KEY: coding sequence
LOCATION: 1 to 759
IDENTIFICATION METHOD: by experiment
NAME/KEY: signal peptide
LOCATION: 1 to 81
IDENTIFICATION METHOD: by experiment
NAME/KEY: mature peptide
LOCATION: 82 to 759
IDENTIFICATION METHOD: by experiment
US-08-974-196-4

Query Match 45.1%; Score 394; DB 2; Length 759;
Best Local Similarity 100.0%; Pred. No. 7.1e-92;
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCAGGATATGTTCACTATGAAGATACTGACCGCAACGAGTCACTGAGGCTTG 60

```

Db	366	CTCCAGCATATGTCACACTATGAAGAATCACTGCACCGGCAACCGACTCATTGGCCCTTG	425
Qy	61	CCGTGCATCTCTTCCCACACGCTGGTAATTGTAACGTGAGAGAGAACCTCCGCAATACTTAT	120
Db	426	CCGTGCATCTCTTCCCACACGCTGGTAATTGTAACGTGAGAGAGAACTCCGCAATACTTAT	485
Qy	121	CTATGAGAGCTGCCGGGGCAATAAACAAGCTACCGCTCTGAGAGGCTCTGCATCTCCG	180
Db	486	CTATGAGAGCTGCCGGGGCAATAAACAAGCTACCGCTCTGAGAGGCTCTGCATCTCCG	545
Qy	181	CTGCTTCCCGCACGACGAGAAATCTCCCCCTGCCCCCTTGCTGAAGSTGGTTCTGGC	240
Db	546	CTGCTTCCCGCACGAGAGAAATCTCCCCCTGCCCCCTTGCTGAAGSTGGTTCTGGC	605
Qy	241	GGGGCGTGTGGTAGTAGGTGTATCTCTCTCCGGAGGCTCOATGCTACTACTATCCG	300
Db	606	GGGGCTGTGTGGTAGTAGGTGTATCTCTCTCCGGAGGCTCTCOATGCTACTACTATCCG	665
Qy	301	GGTGGCACAGSAGAACTCAGAGACGTGGCCCTGCSGACCGTCTTGAAGCTCCGAGATGACAA	360
Db	666	GGTGGCACAGSAGAACTCAGAGACGTGGCCCTGCSGACCGTCTTGAAGCTCCGAGATGACAA	725
Qy	361	GAGGAGCTGGTGAAGAACAATATGTCCTGTGA	394
Db	726	GAGGAGCTGGTGAAGAACAATATGTCCTGTGA	759

RESULT 5
US-09-013-896A-13
; Sequence 13, Application US/09013896A
; Patent No. 6262233

APPLICANT: GENITZ, REIMER
 TITLE OF INVENTION: TISSUE FACTOR PATHWAY INHIBITOR-3
 NUMBER OF SEQUENCES: 31
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: STRINE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
 STREET: 1100 NEW YORK AVE., NW, STE. 600
 CITY: WASHINGTON
 STATE: DC
 COUNTRY: US
 ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/013,896A
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: STEFFE, ERIC K.
 REGISTRATION NUMBER: 36,688
 REFERENCE/DOCKET NUMBER: 1488,1290001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 309-8504
 TELEFAX: (301) 309-8439
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 287 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 US-09-013-896A-13

Query Match	32.3%	Score 282.4;	DB 4,	Length 287;
Best Local Similarity	98.6%;	Pred. No. 1.9e-63;		
Matches 283;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;
46 AGTCACGTGGGCTTCCGTCGATCCTCCCAAGCGTGAAGTTCGATGAGTGGAGAGAACTC 105				

Db	1	AGTCACTGGAGCTTGCCTGGCATCTTCTCCCAACCGTGTACTTTGACGTGGAGGAACTTC	60
Qy	106	CTGCAGTAATCACTCATGTATGAGAGGCTCCGGGGCAATTAAGAAAGCTACCGCTCGAAGA	165
Db	61	CTGCAGTAATCACTCATGTATGAGAGGCTCCGGGGCAATTAAGAAAGCTACCGCTCGAAGA	120
Qy	166	GGCTGCATGCTCCGCTGCTTCCGACAGCAGAGAAATCCTCCCTGCCCCCTTGCTCAA	225
Db	121	GGCTGCATGCTCCGCTGCTTCCGACAGCAGAGAAATCCTCCCTGCCCCCTTGCTCAA	180
Qy	226	GGTGTGTCTTCGACGGGCTTCCTCGAATGAGTGTGATTCCTTCTCGGAGACTCCAT	285
Db	181	GATGTGTCTTCGACGGGCTTCCTCGAATGAGTGTGATTCCTTCTCGGAGACTCCAT	240
Qy	286	GGCTACCTGATCCGGGTGGACAGAGAAACAGAGAGCTCCCTGC	332
Db	241	GGCTACCTTATCCGGGTGGACAGAGAAACAGAGAGCTCCCTGC	287

RESULT 6
US-09-013-896A-14/C
; Sequence 14, Application US/09013896A
Patent No. 7,663,333

```

1      APPLICANT: GENENTZ, REINER
2      TITLE OF INVENTION: TISSUE FACTOR PATHWAY INHIBITOR-3
3      NUMBER OF SEQUENCES: 31
4      CROSS-REFERENCE ADDRESS:
5      ADDRESSEE: STRIMM, KESSLER, GOLDSTEIN & FOX P.L.L.C.
6      STREET: 1100 NEW YORK AVE., NW, STE. 600
7      CITY: WASHINGTON
8      STATE: DC
9      COUNTRY: US
10     ZIP: 20005
11     COMPUTER READABLE FORM:
12     MEDIUM TYPE: Floppy disk
13     COMPUTER: IBM PC compatible
14     OPERATING SYSTEM: PC-DOS/MS-DOS
15     SOFTWARE: PatentIn Release #1.0, Version #1.30
16     CURRENT APPLICATION DATA:
17     APPLICATION NUMBER: US/09/013,896A
18     FILING DATE:
19     CLASSIFICATION: 435
20     ATTORNEY/AGENT INFORMATION:
21     NAME: STEFFE, ERIC K.
22     REGISTRATION NUMBER: 36,688
23     REFERENCE/DOCKET NUMBER: 1488.1290001
24     TELECOMMUNICATION INFORMATION:
25     TELEPHONE: (301) 309-8504
26     TELEFAX: (301) 309-8439
27     INFORMATION FOR SEQ ID NO: 14:
28     SEQUENCE CHARACTERISTICS:
29     LENGTH: 273 base pairs
30     TYPE: nucleic acid
31     STRANDEDNESS: single
32     TOPOLOGY: linear
33     MOLECULE TYPE: DNA (genomic)
34     US-09-013-896A-14

```

Query Match	29.8%;	Score 260;	DB 4;	Length 273;
Best Local Similarity	97.0%;	Pred. No. 1e-57;		
Matches 260; conservative	0;	Mismatches	8;	Indels 0; Gaps 0;

Oy 571 GGTCTCTCTGATGGCTGAGCTGAGAGAGCCAGATGTTCTGCTGATCG 630
 Db 273 GGCTCTCTGATGGCTGAGCTGAGAGAGCCAGATGTTCTGCTGATCG 21.4
 Oy 631 ATTTCTTCCCTCCAGGTAAAGTTTCTTGTAATGATTCATGCTCTTTCTC 690
 Db 213 ATTTCTTCCCTCCAGGTAAAGTTTCTTGTAATGATTCATGCTCTTTCTC 154
 Oy 691 ATCACAAGTGAATGTGAATCTTTCTTTGTGTGATATTATGATTTTTTAAGT 750

Db 153 ATCAGAGAGTATGTGGAAATCGTTCTTTGTTCTGATTATGTTTTTAAAGT 94
 Qy 751 ATAAACAAAAGTTTTTATTTAGCATCTGGAAGAGAAATTAATCAAGTTAAT 810
 Db 93 ATAAACAAAAGTTTTTATTTNNATCTCGAAAGAAAGAAATGTAACAGTTAAT 34
 Qy 811 AAAAAGGGGCTTCCCTTTAGAAATAAA 838
 Db 33 ANAAGGGGCTTCCCTTTAGAAATAAA 6

RESULT 7 US-09-013-896A-15/C

; Sequence 15, Application US/09013896A
 ; Patent No. 6262233
 ; GENERAL INFORMATION:
 ; APPLICANT: GENTZ, REINER
 ; TITLE OF INVENTION: TISSUE FACTOR PATHWAY INHIBITOR-3
 ; NUMBER OF SEQUENCES: 31
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
 ; STREET: 1100 NEW YORK AVE., NW, STE. 600
 ; CITY: WASHINGTON
 ; STATE: DC
 ; COUNTRY: US
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/013,896A
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: STEFFE, ERIC K.
 ; REGISTRATION NUMBER: 36,688
 ; REFERENCE/DOCKET NUMBER: 1488.1290001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (301) 309-8504
 ; TELEFAX: (301) 309-8439
 ; INFORMATION FOR SEQ ID NO: 15:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 256 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; US-09-013-896A-15

Query Match 28.6%; Score 250; DB 4; Length 256;
 Best Local Similarity 99.6%; Pred. No. 3, 6e-55;
 Matches 250; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 588 CCTGACGTCGAGAGAGCCCGAGTGTCTCTCCGATGATGATTTCTTCCAGGT 647
 Db 256 CTTCACTCGAGAGAGCCCGAGTGTCTCTCCGATGATGATTTCTTCCAGGT 197
 Qy 648 AGATTTCTTTGCTTATGTAATTCATTCGCTTTTCTCATCAGAGATGTT 707
 Db 196 AGATTTCTTTGCTTATGTAATTCATTCGCTTTTCTCATCAGAGATGTT 137
 Qy 708 GGAATGCTTTCTTTGTTGTCTGATTTATGTTTTTTAAGTATAAACAAAGTTTTT 767
 Db 136 GGAATGCTTTCTTTGTTGTCTGATTTATGTTTTTTAAGTATAAACAAAGTTTTT 77
 Qy 768 ATTAGCATCTGAAAGAGAAAGTAAATGTAAGTTATATAAAAGGGCTTCCCG 827
 Db 76 ATTAGCATCTGAAAGAGAAAGTAAATGTAAGTTATATAAAAGGGCTTCCCG 17
 Qy 828 TTAGAAATAAA 838

Db 16 TTAGAAATAAA 6

RESULT 8 US-09-013-896A-17

; Sequence 17, Application US/09013896A
 ; Patent No. 6262233
 ; GENERAL INFORMATION:
 ; APPLICANT: GENTZ, REINER
 ; TITLE OF INVENTION: TISSUE FACTOR PATHWAY INHIBITOR-3
 ; NUMBER OF SEQUENCES: 31
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
 ; STREET: 1100 NEW YORK AVE., NW, STE. 600
 ; CITY: WASHINGTON
 ; STATE: DC
 ; COUNTRY: US
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/013,896A
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: STEFFE, ERIC K.
 ; REGISTRATION NUMBER: 36,688
 ; REFERENCE/DOCKET NUMBER: 1488.1290001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (301) 309-8504
 ; TELEFAX: (301) 309-8439
 ; INFORMATION FOR SEQ ID NO: 17:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 201 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; US-09-013-896A-17

Query Match 21.5%; Score 188; DB 4; Length 201;
 Best Local Similarity 99.0%; Pred. No. 2, 5e-39;
 Matches 199; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 310 GAGGAACGAGAGCGTGCCTGCGCACCGTCTGAGAGCTCCGAGATGACAAGAGAGCT 369
 Db 1 GAGGAACGAGAGCGTGCCTGCGCACCGTCTGAGAGCTCCGAGATGACAAGAGAGCT 60
 Qy 370 GGTGAAGAACACATATGTCTGTGACCGCCCTGTGCGCAAGAGACT-GGGAAGGAGGG 428
 Db 61 GGTGAAGAACACATATGTCTGTGACCGCCCTGTGCGCAAGAGACTGGGAAGGAGGG 120
 Qy 429 GAGACTATGTGACCTTTTATTAATAGAGGATTTGACTGGATTTGATGATCTTGG 488
 Db 121 GAGACTATGTGACCTTTTATTAATAGAGGATTTGACTGGATTTGATGATCTTGG 180
 Qy 489 GGCTGAGGTCTGTTCTCTGG 509
 Db 181 GGCTGAGGTCTGTTCTCTGG 201

RESULT 9 US-09-404-879A-169

; Sequence 169, Application US/09404879A
 ; Patent No. 6468546
 ; GENERAL INFORMATION:
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: King, Gordon E.
 ; APPLICANT: Algate, Paul A.

```

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404,879A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 169
; LENGTH: 276
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-404-879A-169

```

```

Query Match          19.2%; Score 167.4; DB 4; Length 276;
Best Local Similarity 99.4%; Pred. No. 5.6e-34;
Matches 168; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 CTCACGATATGTTCAATATGAAAGATCTGACCCGCAACGAGTCACTGGGCTTG 60
DB 88 CTCACGATATGTTCAATATGAAAGATCTGACCCGCAACGAGTCACTGGGCTTG 147
QY 61 CCGTGATCCTTCCACGCTGTGACTTTGACGTGAGAGAACTCTGCAATACTTCAT 120
DB 148 CCGTGATCCTTCCACGCTGTGACTTTGACGTGAGAGAACTCTGCAATACTTCAT 207
QY 121 CTATGAGGCTGCGGGGCAATAGAAACAGCTACCGCTCTGAGAGAGCC 169
DB 208 CTATGAGGCTGCGGGGCAATAGAAACAGCTACCGCTCTGAGAGAGAC 256

```

```

RESULT 10
US-09-404-879A-356
; Sequence 356, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404,879A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 356
; LENGTH: 207
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-404-879A-356

```

```

Query Match          19.1%; Score 167; DB 4; Length 207;
Best Local Similarity 100.0%; Pred. No. 6.2e-34;
Matches 167; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 CTCACGATATGTTCAATATGAAAGATCTGACCCGCAACGAGTCACTGGGCTTG 60
DB 41 CTCACGATATGTTCAATATGAAAGATCTGACCCGCAACGAGTCACTGGGCTTG 100
QY 61 CCGTGATCCTTCCACGCTGTGACTTTGACGTGAGAGAACTCTGCAATACTTCAT 120
DB 101 CCGTGATCCTTCCACGCTGTGACTTTGACGTGAGAGAACTCTGCAATACTTCAT 160
QY 121 CTATGAGGCTGCGGGGCAATAGAAACAGCTACCGCTCTGAGAGAG 167
DB 161 CTATGAGGCTGCGGGGCAATAGAAACAGCTACCGCTCTGAGAGAG 207

```

```

RESULT 11
US-09-404-879A-365/c
; Sequence 365, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:

```

```

; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404,879A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 365
; LENGTH: 371
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-404-879A-365

```

```

Query Match          19.1%; Score 167; DB 4; Length 371;
Best Local Similarity 100.0%; Pred. No. 8.1e-34;
Matches 167; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 CTCACGATATGTTCAATATGAAAGATCTGACCCGCAACGAGTCACTGGGCTTG 60
DB 167 CTCACGATATGTTCAATATGAAAGATCTGACCCGCAACGAGTCACTGGGCTTG 108
QY 61 CCGTGATCCTTCCACGCTGTGACTTTGACGTGAGAGAACTCTGCAATACTTCAT 120
DB 107 CCGTGATCCTTCCACGCTGTGACTTTGACGTGAGAGAACTCTGCAATACTTCAT 48
QY 121 CTATGAGGCTGCGGGGCAATAGAAACAGCTACCGCTCTGAGAGAG 167
DB 47 CTATGAGGCTGCGGGGCAATAGAAACAGCTACCGCTCTGAGAGAG 1

```

```

RESULT 12
US-09-404-879A-168/c
; Sequence 168, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404,879A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 168
; LENGTH: 276
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (1)-(276)
; OTHER INFORMATION: n = A,T,C or G
US-09-404-879A-168

```

```

Query Match          18.9%; Score 165.4; DB 4; Length 276;
Best Local Similarity 98.2%; Pred. No. 1.8e-33;
Matches 166; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 CTCACGATATGTTCAATATGAAAGATCTGACCCGCAACGAGTCACTGGGCTTG 60
DB 189 CTCACGATATGTTCAATATGAAAGATCTGACCCGCAACGAGTCACTGGGCTTG 130
QY 61 CCGTGATCCTTCCACGCTGTGACTTTGACGTGAGAGAACTCTGCAATACTTCAT 120
DB 129 CCGTGATCCTTCCACGCTGTGACTTTGACGTGAGAGAACTCTGCAATACTTCAT 70
QY 121 CTATGAGGCTGCGGGGCAATAGAAACAGCTACCGCTCTGAGAGAGCC 169
DB 69 CTATGAGGCTGCGGGGCAATAGAAACAGCTACCGCTCTGAGAGAGAC 21

```


SEQUENCE CHARACTERISTICS:
; LENGTH: 783 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-030-607-15

Query Match 8.3%; Score 72.8; DB 4; Length 783;
Best Local Similarity 61.7%; Pred. No. 1.6e-09;
Matches 113; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 23 GAAGAACTACTGACCGCCAGCAGTCACTGGGCTTGGCGTGCATCTCTTCCACGCTGG 82
DB 252 GAAGACTACTGCTGCTGCATCCACANGTGGGTGCTGCCGGGCTCTTCCACGCTGG 311
QY 83 TACTTGAAGTGGAGAGAACTCTTGCATTAATTGATCTATGAGAGCTGCCGGGGCAAT 142
DB 312 TACTATGACCCCAAGAGCAGATCTGCAGAGATTGCTTATGAGAGCTGCTTGGCAAC 371
QY 143 AAGACAGCTACCGCTCTGAGAGGCTGCATGCTCCGCTTCCGACAGCAGAGAAAT 202
DB 372 AAGAACACTACTCTTGGGAGAGAGTGCATTCTANCTGTONGGGTGTGCAGAGTGGG 431
QY 203 CCT 205
DB 432 CCT 434

Search completed: January 10, 2003, 03:22:17
Job time : 56.01 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 10, 2003, 02:45:21 (Search time 199.668 Seconds)

(without alignments)
1924.247 Million cell updates/sec

Title: US-09-825-682a-56

Perfect score: 873

Sequence: 1 cccacgcatgttccact.....tcaaaaaaaaaaaaaa 873

Scoring table:

IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 389086 segs, 220051671 residues

Total number of hits satisfying chosen parameters: 778172

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCM_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PT06_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	862	98.7	1610	US-09-827-948-1	Sequence 1, Appl
2	843	96.6	1527	US-09-822-830A-95	Sequence 95, Appl
3	825.2	94.5	1544	US-09-880-107-3429	Sequence 3429, Ap
4	758.2	86.8	1649	US-09-925-301-424	Sequence 424, App
5	316.4	36.2	320	US-09-604-287A-293	Sequence 293, App
6	316.4	36.2	320	US-09-339-338-293	Sequence 293, App
7	316.4	36.2	320	US-10-007-805-293	Sequence 293, App
8	312.2	35.8	331	US-09-869-708-166	Sequence 166, App
9	312.2	35.8	331	US-09-880-107-2048	Sequence 2048, Ap
10	288.4	33.0	320	US-09-867-701-510	Sequence 510, App
11	282.4	32.3	287	US-09-827-948-13	Sequence 13, Appl
12	273	31.3	285	US-09-815-343-703	Sequence 32, Appl
13	262	30.0	286	US-09-815-343-703	Sequence 703, Appl
14	260	29.8	273	US-09-827-948-14	Sequence 14, Appl
15	253.8	29.1	287	US-09-815-343-1123	Sequence 15, Appl
16	250	28.6	256	US-09-827-948-15	Sequence 15, Appl
17	238.4	27.3	241	US-09-604-287A-376	Sequence 376, App
18	238.4	27.3	241	US-10-007-805-376	Sequence 376, App
19	226.4	25.9	287	US-09-815-343-1258	Sequence 1258, App

20	188	21.5	201	US-09-827-948-17	Sequence 17, Appl
21	186.2	21.3	396	US-09-970-966-18	Sequence 18, Appl
22	186.2	21.3	396	US-09-825-294-18	Sequence 18, Appl
23	167.4	19.2	276	US-09-884-441-169	Sequence 169, App
24	167	19.1	207	US-09-884-441-356	Sequence 356, App
25	167	19.1	224	US-09-815-343-1343	Sequence 1343, Ap
26	167	19.1	234	US-09-815-343-1063	Sequence 1063, Ap
27	167	19.1	234	US-09-815-343-1129	Sequence 1129, Ap
28	167	19.1	234	US-09-815-343-1191	Sequence 1191, Ap
29	167	19.1	371	US-09-884-441-355	Sequence 365, App
30	165.4	18.9	276	US-09-815-343-1190	Sequence 1190, Ap
31	165.4	18.9	234	US-09-884-441-168	Sequence 168, App
32	150.6	17.3	302	US-09-815-343-991	Sequence 991, App
33	146.8	16.8	199	US-09-604-287A-125	Sequence 125, App
34	146.8	16.8	199	US-09-339-338-125	Sequence 125, App
35	146.8	16.8	199	US-10-007-805-125	Sequence 125, App
36	99.8	11.4	396	US-09-970-966-75	Sequence 75, Appl
37	99.8	11.4	396	US-09-825-294-75	Sequence 75, Appl
38	96.6	11.1	425	US-09-960-352-6197	Sequence 6197, Ap
39	95.4	10.9	369	US-09-960-352-12098	Sequence 12098, A
40	93.8	10.7	429	US-09-960-352-4987	Sequence 4987, Ap
41	72.8	8.3	783	US-09-232-880-15	Sequence 15, Appl
42	72.8	8.3	783	US-10-012-896-15	Sequence 15, Appl
43	72.8	8.3	783	US-09-895-793-15	Sequence 15, Appl
44	72.8	8.3	783	US-09-895-814-15	Sequence 15, Appl
45	72.8	8.3	783	US-09-759-143-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-09-827-948-1
Sequence 1, Application US/09827948
Patent No. US20010029034A1
GENERAL INFORMATION:
APPLICANT: Gentz, Reiner, L.
APPLICANT: Hsu, Tzu-An
APPLICANT: Rosen, Craig A.
APPLICANT: Ni, Jian
TITLE OF INVENTION: Tissue Factor Pathway Inhibitor-3
FILE REFERENCE: 1488.1290002
CURRENT APPLICATION NUMBER: US/09/827, 948
CURRENT FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: US 09/013, 896
PRIOR FILING DATE: 1998-01-27
NUMBER OF SEQ ID NOS: 31
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 1610
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (361)..(1116)
NAME/KEY: sig.peptide
LOCATION: (361)..(439)
NAME/KEY: mat.peptide
LOCATION: (442)..(1116)
US-09-827-948-1

Query Match 98.7%; Score 862; DB 10; Length 1610;
Best Local Similarity 99.9%; Pred. No. 1.6e-203;
Matches 873; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY	1	CCCCAGCATATGTTCACTAATGAAGAACTATGACCGCAACGACGACGCGGCGCTTG	60
DB	726	CTCCAGCGATATGTTCACTAATGAAGAACTATGACCGCAACGACGACGCGGCGCTTG	785
QY	61	CCGTGATCCTTCCACGCGTGAATCTTGAAGTGAAGAGAACTCTGCAATTAATCTCAT	120
DB	786	CCGTGATCCTTCCACGCGTGAATCTTGAAGTGAAGAGAACTCTGCAATTAATCTCAT	845

Db	Seq	Score	Length	Indels	Gaps
Db	1206	99.58	1544	9	1265
Qy	480	94.58	1544	9	539
Qy	540	94.58	1544	9	599
Db	1146	94.58	1544	9	1205
Qy	420	94.58	1544	9	479
Db	1086	94.58	1544	9	1145
Qy	361	94.58	1544	9	419
Db	1026	94.58	1544	9	1085
Qy	301	94.58	1544	9	360
Db	966	94.58	1544	9	1025
Qy	241	94.58	1544	9	300
Db	906	94.58	1544	9	965
Qy	181	94.58	1544	9	240
Db	846	94.58	1544	9	905
Qy	786	94.58	1544	9	845
Db	726	94.58	1544	9	785
Qy	666	94.58	1544	9	725
Db	606	94.58	1544	9	605
Qy	546	94.58	1544	9	545
Db	486	94.58	1544	9	485
Qy	426	94.58	1544	9	425
Db	366	94.58	1544	9	365
Qy	306	94.58	1544	9	305
Db	246	94.58	1544	9	245
Qy	186	94.58	1544	9	185
Db	126	94.58	1544	9	125
Qy	66	94.58	1544	9	65
Db	6	94.58	1544	9	5

```

QY 600 CAGAGCCCGCAGTGTTCCTCAGCTGATGATTTCTTCTCCACAGTAGAGTTTCTTT 659
Db 1266 CAGCAGCCCGCAGTGTTCCTCAGCTGATGATTTCTTCTCCACAGTAGAGTTTCTTT 1325
QY 660 GCTATGTTGAATTCATGTCCTC--TTTCTCATCACAAGAGTATGTTGAATGCTTT 717
Db 1326 GCTATGTTGAATTCATGTCCTCCTTTCTCCNATCACAAGATGATGTGGAAATGCTTT 1385
QY 718 CTTTGTGTTGCTGATTTATGATTTTAAATGATTAACAAAAGTTTATTTATGACATTC 777
Db 1386 CTTTGTGTTGCTGATTTATGATTTTAAATGATTAACAAAAGTTTATTTATGACATTC 1445
QY 778 TGAAGAAGAAATGAATAATGATCAAGTTATTAAGAGGAGGCTTCCCTTAGAATAA 837
Db 1446 TGAAGAAGAAATGAATAATGATCAAGTTATTAAGAGGAGGCTTCCCTTAGAATAA 1505
QY 838 ATTT-CAGCATGTGCTTTCAAAAAAAAAAAAAAAAAAAAA 873
Db 1506 ATTTCCAGCATGTGCTTTCAAAAAAAAAAAAAAAAAAAAA 1542

RESULT 4
US-09-925-301-424
; Sequence 424, Application US/0925301
; Patent No. US0020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925.301
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05862
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 424
; LENGTH: 1649
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-301-424

Query Match 86.8%; Score 758.2; DB 10; Length 1649;
Best Local Similarity 95.6%; Pred. No. 7.2e-178;
Matches 833; Conservative 0; Mismatches 33; Indels 5; Gaps 5

QY 1 CTCACAGCATATGTTCAACTATGAAAGATACTGCACCGCCACGACGATCATCGGACCTTG 60
Db 784 CTCACAGCATATGTTCAACTATGAAAGATACTGCACCGCCACGACGATCATCGGACCTTG 843
QY 61 CCGTGATCTTTCCACAGCTGTGACTTTGACGTGGAAGAACTCCTGGAATAACTTCAT 120
Db 844 CCGTGATCTTTCCACAGCTGTGACTTTGACGTGGAAGAACTCCTGGAATAACTTCAT 903
QY 121 CTAATGAGAGCTCCGAGGAGCAATAGAAACAGCTACCGCTCGAGAGAGCTCGAAGTCTCG 180
Db 904 CTAATGAGAGCTCCGAGGAGCAATAGAAACAGCTACCGCTCGAGAGAGCTCGAAGTCTCG 963
QY 181 CTGCTTCGCGACGAGAGAAATCTCTCCCTGCTGCTTGGCTCAAAAGTGTGTGTTCTGCG 240
Db 964 CTGCTTCGCGACGAGAGAAATCTCTCCCTGCTGCTTGGCTCAAAAGTGTGTGTTCTGCG 1023
QY 241 GGGAGCTTTCCGTAAGTGTGATCTCTCTCCGAGAGAGCTCCCATGATCTACCTGATCCG 300
Db 1024 GGGAGCTTTCCGTAAGTGTGATCTCTCTCCGAGAGAGCTCCCATGATCTACCTGATCCG 1083
QY 301 GGTGAGACGAGAGAAACGAGAGAGTGCCTCTGCGACCGCTCGAGAGCTCCGAGATGACAA 360
Db 1084 GGTGAGACGAGAGAAACGAGAGAGTGCCTCTGCGACCGCTCGAGAGCTCCGAGATGACAA 1143
QY 361 GAGACAGCTGTGAAAGAACATATGTCTCTGTGACCGCCTGTGCGCAAGAGAACTGGGA 420

```

```

Db 1144 GAGACAGCTGGAGAAACATATGTCCTG-CCGCCCTGTCGCAAGAGACTG95G 1202
|
|
|
Qy 421 AAGGAGGAGGAGACATGCTGTA-CTTTTAAATAGAGGATGACTCGATTTAGTG 480
|
|
|
Db 1203 AAGGAGGAGGAGACATGCTGTA-CTTTTAAATAGAGGATGACTCGATTTAGTG 1261
|
|
|
Qy 481 ATCATTAAGGCTGAGGCTGTTCTCTGAGAGGTAGAGCGCTGCTTCTGCTG3CAG 540
|
|
|
Db 1262 ATCATTAAGGCTGAGGCTGTTCTCTGAGAGGTAGAGCGCTGCTTCTGCTG3CAG 1321
|
|
|
Qy 541 GATGGGTTTGTCTTGAGAAATCCTCTAGAGAGGCTCTCTCGACATGSCCTGAGCT95C 600
|
|
|
Db 1322 GATGGGTTTGTCTTGAGAAATCCTCTAGAGAGGCTCTCTCGACATGSCCTGAGCT95C 1381
|
|
|
Qy 601 AGCAGCCCGAGTGTCTCTGCTGATGATTTCTTCTCTCGAGTGAAGTTTCTTG 660
|
|
|
Db 1382 AGCAGCCCGAGTGTCTCTGCTGATGATTTCTTCTCTCGAGTGAAGTTTCTTG 1440
|
|
|
Qy 661 CTATGTTGAATTCATGCTCTCTTCTCTCATACAGAGTGTGAAATGCTTCTT 720
|
|
|
Db 1441 CTATGTTGAATTCATGCTCTCTTCTCTCATACAGAGTGTGAAATGCTTCTT 1499
|
|
|
Qy 721 TTGTTGCTGATTTATAGCTTTTAAATTAACAAAGTTTATTAAGATTTCTGA 780
|
|
|
Db 1500 TTGTTGCTGATTTATAGCTTTTAAATTAACAAAGTTTATTAAGATTTCTGA 1559
|
|
|
Qy 781 AAGAGAGAAATGAATGATACAGTTTAATTAAGAGGAGGCTCTCTTGAATTAAT 840
|
|
|
Db 1560 AAGAGAGAAATGAATGATACAGTTTAATTAAGAGGAGGCTCTCTTGAATTAAT 1618
|
|
|
Qy 841 TCAGCATGTGCTTCACAAAAA871
|
|
|
Db 1619 AAAAAA871
|
|
|

```

```

RESULT 5
US-09-604-287A-293
; Sequence 293, Application US/09604287A
; Patent No. US20020064872A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Micham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C7
; CURRENT APPLICATION NUMBER: US/09/604,287A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 293
; LENGTH: 320
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-604-287A-293

```

```

Query Match 36.2%; Score 316.4; DB 10; Length 320;
Best Local Similarity 99.7%; Pred. No. 4,6e-69;
Matches 317; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 519 CGGCTGCTTCTGCTGAGGAGGATGAGGCTTTGTTGAATCCTTAGAGAGCTCCTC 578
|
|
|
Db 2 CGGCTGCTTCTGCTGAGGAGGATGAGGCTTTGTTGAATCCTTAGAGAGCTCCTC 61
|
|
|
Qy 579 CTGCAATGCTGCTGAGTCTGAGCAGACCCCGAGTGTCTCTGCTGATGATTTCTT 638
|
|
|
Db 62 CTGCAATGCTGCTGAGTCTGAGCAGACCCCGAGTGTCTCTGCTGATGATTTCTT 121
|
|
|
Qy 639 CTGCAATGAGTGTCTTCTTGTATGTTGAATTCATGCTCTTTCTCATCAGAGA 698
|
|
|

```

```

Db 122 CTCACAGTAGAGTTCTTCTGCTAATGTTGAATTCATGCTCTTTCTCATCAGAGA 181
|
|
|
Qy 699 AGTAGATGAGATGATGCTTTCTTTGTTGCTGATTTATGATTTTAAATTAACAA 758
|
|
|
Db 182 AGTAGATGAGATGATGCTTTCTTTGTTGCTGATTTATGATTTTAAATTAACAA 241
|
|
|
Qy 759 AAGTTTTATTAAGATTCGAAAGAGAAAGTAAATGTACAAAGTTTAATTAAGAG 818
|
|
|
Db 242 AAGTTTTATTAAGATTCGAAAGAGAAAGTAAATGTACAAAGTTTAATTAAGAG 301
|
|
|
Qy 819 GCCTTCCCTTTAGATA 836
|
|
|
Db 302 GCCTTCCCTTTAGATA 319
|
|
|

```

```

RESULT 6
US-09-339-338-293
; Sequence 293, Application US/09339338A
; Patent No. US20020102602A1
; GENERAL INFORMATION:
; APPLICANT: Yugu, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Micham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; FILE REFERENCE: 210121.470C2
; CURRENT APPLICATION NUMBER: US/09/339,338A
; CURRENT FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 315
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 293
; LENGTH: 320
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-339-338-293

```

```

Query Match 36.2%; Score 316.4; DB 10; Length 320;
Best Local Similarity 99.7%; Pred. No. 4,6e-69;
Matches 317; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 519 CGGCTGCTTCTGCTGAGGAGGATGAGGCTTTGTTGAATCCTTAGAGAGCTCCTC 578
|
|
|
Db 2 CGGCTGCTTCTGCTGAGGAGGATGAGGCTTTGTTGAATCCTTAGAGAGCTCCTC 61
|
|
|
Qy 579 CTGCAATGCTGAGTCTGAGCAGACCCCGAGTGTCTCTGCTGATGATTTCTT 638
|
|
|
Db 62 CTGCAATGCTGAGTCTGAGCAGACCCCGAGTGTCTCTGCTGATGATTTCTT 121
|
|
|
Qy 639 CTGCAATGCTGAGTCTGAGCAGACCCCGAGTGTCTCTGCTGATGATTTCTT 698
|
|
|
Db 122 CTGCAATGCTGAGTCTGAGCAGACCCCGAGTGTCTCTGCTGATGATTTCTT 181
|
|
|
Qy 699 AGTAGATGAGATGATGCTTTCTTTGTTGCTGATTTATGATTTTAAATTAACAA 758
|
|
|
Db 182 AGTAGATGAGATGATGCTTTCTTTGTTGCTGATTTATGATTTTAAATTAACAA 241
|
|
|
Qy 759 AAGTTTTATTAAGATTCGAAAGAGAAAGTAAATGTACAAAGTTTAATTAAGAG 818
|
|
|
Db 242 AAGTTTTATTAAGATTCGAAAGAGAAAGTAAATGTACAAAGTTTAATTAAGAG 301
|
|
|
Qy 819 GCCTTCCCTTTAGATA 836
|
|
|
Db 302 GCCTTCCCTTTAGATA 319
|
|
|

```

```

RESULT 7
US-10-007-805-293
; Sequence 293, Application US/10007805
; Patent No. US20020150581A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu

```

```

1 RESULT 8
2 US-09-969-708-166/c
3 : Sequence 166, Application US/03969708
4 : Patent No. US20020102512A1
5 : GENERAL INFORMATION:
6 : APPLICANT: Augustus, Meena
7 : TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
8 : TITLE OF INVENTION: Sets
9 : FILE REFERENCE: 6698290-70
10 : CURRENT APPLICATION NUMBER: US/09/969,708
11 : CURRENT FILING DATE: 2001-10-03
12 : PRIOR APPLICATION NUMBER: US/60/237,606
13 : PRIOR FILING DATE: 2000-10-03
14 : PRIOR APPLICATION NUMBER: US/60/237,608
15 : PRIOR FILING DATE: 2000-10-03
16 : PRIOR APPLICATION NUMBER: US/60/237,425
17 : PRIOR FILING DATE: 2000-10-03
18 : NUMBER OF SEQ ID NOS: 658
19 : SOFTWARE: PatentIn version 3.0
20 : SEQ ID NO 166
21 : LENGTH: 331

```

	Query Match	35.8%	Score 312.2	DB 10	Length 331
	Best Local Similarity	98.2%	Pred. 31.2	5.1e-68	
	Matches 325	Conservative 0	Mismatches 5	Indels 1	Gaps 1
Oy	484	ATTAGGCGCTGAGGCTCTTCTCTGAGGATGAGACGCGTG	-CTTCTGAGTCTGACAGG	542	
Ob	331	ATTAGGCGCTGAGGCTCTTCTCTGAGGATGAGACGCGTG	-CTTCTGAGTCTGACAGG	272	

QY 543 ATGGTTCTTGGAAATCTCTAGAGAGCTCTCTCGATGCGCTGAGTCTGGCAG 602
 DB 271 ATGGTTCTTGGAAATCTCTAGAGAGCTCTCTCGATGCGCTGAGTCTGGCAG 212
 QY 603 CAGCCCGAGTGTCTCTCGATGAGATTTCTTCCCGAGGATGATTTCTTGTCT 662
 DB 211 CAGCCCGAGTGTCTCTCGATGAGATTTCTTCCCGAGGATGATTTCTTGTCT 152
 QY 663 TATGTGAATTCATTCGCTCTTCTTCTCAACAGAGTATGTTGAATCGTTCTTT 722
 DB 151 TATGTGAATTCATTCGCTCTTCTTCTCAACAGAGTATGTTGAATCGTTCTTT 92
 QY 723 GTTGTCTGATTTATGTTTTTTTAACTATAACAAAGTTTTTTTATGACATTCGAAA 782
 DB 91 GTTGTCTGATTTATGTTTTTTTAACTATAACAAAGTTTTTTTATGACATTCGAAA 32
 QY 783 GAAGGAAGTAAATGTTCAGTTTATATA 813
 DB 31 GAAGGAAGTAAATGTTCAGTTTATATA 1

RESULT 10

US-09-867-701-530
 ; Sequence 530, Application US/09867701
 ; Patent No. US20020132237A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Aglate, Paul A.
 ; APPLICANT: Jones, Robert
 ; APPLICANT: Hatziocker, Susan L.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; FILE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
 ; FILE REFERENCE: 210121.497
 ; CURRENT APPLICATION NUMBER: US/09/867,701
 ; CURRENT FILING DATE: 2001-05-29
 ; NUMBER OF SEQ ID NOS: 10912
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 530
 ; LENGTH: 320
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(320)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-867-701-530

Query Match 33.0%; Score 288.4; DB 10; Length 320;

Best Local Similarity 99.3%; Pred. No. 3; 8e-62; Mismatches 1; Indels 1; Gaps 1;

QY 533 TCTGGCAGGAGTGGGTTGCTTTGGAATCTCTAGAGAGCTCTCTCGATGCGCTGAG 592
 DB 1 TCTGGCAGGAGTGGGTTGCTTTGGAATCTCTAGAGAGCTCTCTCGATGCGCTGAG 60
 QY 593 AGTGTGACAGACCCCGAGTGTCTCTCGATGAGATTTCTTCTCCAGGTAGAGT 652
 DB 61 AGTGTGACAGACCCCGAGTGTCTCTCGATGAGATTTCTTCTCCAGGTAGAGT 119
 QY 653 TTTCTTGTCTTATGTTGAATTCATTCGCTCTTCTCTCAACAGAGTATGTTGAAT 712
 DB 120 TTTCTTGTCTTATGTTGAATTCATTCGCTCTTCTCTCAACAGAGTATGTTGAAT 179
 QY 713 CGTTCTTGTCTTATGTTGAATTTGTTTTTAACTATAACAAAGTTTTTATGATG 772
 DB 180 CGTTCTTGTCTTATGTTGAATTTGTTTTTAACTATAACAAAGTTTTTATGATG 239
 QY 773 CATTCGAAGAAGAAAGTAAATGTACAGATTTATATAAAGGAGGCTTCCCTTTAG 832
 DB 240 CATTCGAAGAAGAAAGTAAATGTACAGATTTATATAAAGGAGGCTTCCCTTTAG 299
 QY 833 AA 834

DB 300 GA 301

RESULT 11

US-09-827-948-13
 ; Sequence 13, Application US/09827948
 ; Patent No. US20010029034A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gentz, Reiner, L.
 ; APPLICANT: Hsu, Tzu-An
 ; APPLICANT: Rosen, Craig A.
 ; APPLICANT: Ni, Jian
 ; TITLE OF INVENTION: Tissue Factor Pathway Inhibitor-3
 ; FILE REFERENCE: 1488.1290002
 ; CURRENT APPLICATION NUMBER: US/09/827,948
 ; CURRENT FILING DATE: 2001-04-06
 ; PRIOR APPLICATION NUMBER: US 09/013,896
 ; PRIOR FILING DATE: 1998-01-27
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 13
 ; LENGTH: 287
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (182)
 ; OTHER INFORMATION: n is A, C, T, or G
 ; NAME/KEY: misc_feature
 ; LOCATION: (193)
 ; OTHER INFORMATION: n is A, C, T, or G
 ; NAME/KEY: misc_feature
 ; LOCATION: (229)
 ; OTHER INFORMATION: n is A, C, T, or G
 US-09-827-948-13

Query Match 32.3%; Score 282.4; DB 10; Length 287;

Best Local Similarity 98.6%; Pred. No. 1; 1e-60; Mismatches 4; Indels 0; Gaps 0;

QY 46 AGTCACTGGAGCCTTCGCTGATCCTTCCACGCTGTACTTTCAGCGAGAGAACTC 105
 DB 1 AGTCACTGGAGCCTTCGCTGATCCTTCCACGCTGTACTTTCAGCGAGAGAACTC 60
 QY 106 CTGCAATTAATCTTATCTATGAGAGCTGCGCGGCAATTAAGAAACAGCTTCTGAGA 165
 DB 61 CTGCAATTAATCTTATCTATGAGAGCTGCGCGGCAATTAAGAAACAGCTTCTGAGA 120
 QY 166 GGCTGTATGCTCCGCTGCTCCGCGAGAGAGAGATCTCCCTGCCCCCTTGCTTCAA 225
 DB 121 GGCTGTATGCTCCGCTGCTCCGCGAGAGAGAGATCTCCCTGCCCCCTTGCTTCAA 180
 QY 226 GGTGTGTTTCTGGCGGGGCTTTGATGATGATGATCTCTTCTGGAGAGCTTCAT 285
 DB 181 GGTGTGTTTCTGGCGGGGCTTTGATGATGATGATGATGATCTCTTCTGGAGAGCTTCAT 240
 QY 286 GGTCTACCTGATCCGGGTGGACAGAGAAACAGAGAGCGTGCCTGC 332
 DB 241 GGTCTACCTGATCCGGGTGGACAGAGAAACAGAGAGCGTGCCTGC 287

RESULT 12

US-09-815-343-32/C
 ; Sequence 32, Application US/09815343
 ; Patent No. US2001005596A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Weagner, Madeleine
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: King, Gordon E.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
 ; FILE REFERENCE: 210121.504
 ; CURRENT APPLICATION NUMBER: US/09/815,343

GenCore Version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 10, 2003, 01:46:01 : Search time 1874.25 Seconds
(without alignments)
7543.621 Million cell updates/sec

Title: US-09-825-682a-56
Perfect score: 873
Sequence: 1 ctccgcgcgcatacttcaact.....tcaaaaaaaaaaaaaaaaaa 873

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seg length: 0
Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_esthum:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	803.2	92.0	917	9	AL576113
C 2	802.8	92.0	910	9	AL572703
C 3	785	89.9	1001	9	AL548687
C 4	783.4	89.7	923	10	BE621676
C 5	782.4	89.6	1061	9	AL575665
C 6	781.6	89.5	983	9	AL573986

C 7	776.2	88.9	938	9	AL575631	AL575631
C 8	776	88.9	1052	9	AL573346	AL573346
C 9	770.8	88.3	877	9	AL576440	AL576440
C 10	770.6	88.3	950	9	AL576831	AL576831
C 11	752.8	86.2	1062	9	AL575764	AL575764
C 12	748	85.7	964	9	AL560664	AL560664
C 13	747.8	85.7	859	9	AL582333	AL582333
C 14	742.2	85.0	990	9	AL571794	AL571794
C 15	738	84.5	750	14	BM977432	BM977432
C 16	735.6	84.3	985	9	AL551363	AL551363
C 17	735.2	84.2	889	14	BQ439520	BQ439520
C 18	734.6	84.1	810	9	AL188641	AL188641
C 19	734.6	84.1	942	12	BE741988	BE741988
C 20	734	84.1	951	9	AL536057	AL536057
C 21	730.2	83.6	814	9	AL1983859	AL1983859
C 22	716.6	82.1	1038	9	AL571204	AL571204
C 23	700.6	80.3	986	12	BG681770	BG681770
C 24	700.2	80.2	1143	14	BM924750	BM924750
C 25	692.2	79.3	938	14	BQ949537	BQ949537
C 26	689.4	79.0	726	14	BM979926	BM979926
C 27	689	78.9	713	14	BQ181625	BQ181625
C 28	685.4	78.5	728	14	BQ575516	BQ575516
C 29	681	78.0	746	12	BE743960	BE743960
C 30	677	77.5	789	13	BM044691	BM044691
C 31	675	77.3	958	13	B1523756	B1523756
C 32	673	77.1	758	12	BE744503	BE744503
C 33	672	77.0	745	9	AL1819257	AL1819257
C 34	668.4	76.6	949	12	BE741701	BE741701
C 35	667.4	76.4	880	14	BQ227721	BQ227721
C 36	665.8	76.3	716	9	AL574271	AL574271
C 37	663.2	76.0	719	10	BE613272	BE613272
C 38	663	75.9	908	12	BE747343	BE747343
C 39	662.4	75.9	748	10	AM190372	AM190372
C 40	659.4	75.5	703	9	AL1982697	AL1982697
C 41	654.8	75.0	951	14	BQ947368	BQ947368
C 42	649.6	74.4	727	13	B1087012	B1087012
C 43	649.2	74.4	800	9	AL572724	AL572724
C 44	645.4	73.9	842	12	BE740873	BE740873
C 45	644	73.8	668	14	BM978012	BM978012

ALIGNMENTS

RESULT 1
AL576113/c 917 bp mRNA linear EST 16-FEB-2001
LOCUS AL576113 LFI NFL006.PL2 Homo sapiens cDNA clone CSOD1072YC24 3
DEFINITION prime, mRNA sequence.

ACCESSION AL576113
VERSION AL576113 GI:12937934
KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 917)
AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polyes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

1. 917
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSOD1072YC24"
/clone_1ib="LFI_NFL006.PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com>

BASE COUNT 243 a 254 c 216 g 200 t 4 others
ORIGIN

Query Match 92.0%; Score 803.2; DB 9; Length 917;
Best Local Similarity 99.4%; Pred. No. 1.4e-136;
Matches 813; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

```

QY 1 CTCACGATATGTTCAACTATGAAGAATCTGACCGCCCAACGCACTGCTGCGCTTG 60
DB 818 CTCACGATATGTTCAACTATGAAGAATCTGACCGCCCAACGCACTGCTGCGCTTG 759
QY 61 CCGTGATCTCTCCACGCTGCTGCTTGTACGTGAGAGAGAACTCTGCAATTAATTCT 120
DB 758 CCGTGATCTCTCCACGCTGCTGCTTGTACGTGAGAGAGAACTCTGCAATTAATTCT 699
QY 121 CTATGAGAGCTGCGGGGCAATAAGAACACTACCGCTCTGAGAGAGCTGCAATGCTCCG 180
DB 698 CTATGAGAGCTGCGGGGCAATAAGAACACTACCGCTCTGAGAGAGCTGCAATGCTCCG 639
QY 181 CTGCTTCCGACAGAGAGAAATCCCTCCCTGCTTGCTGCTCAAGGTGTGTCTGCG 240
DB 638 CTGCTTCCGACAGAGAGAAATCCCTCCCTGCTTGCTGCTCAAGGTGTGTCTGCG 579
QY 241 GGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
DB 578 GGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 519
QY 301 GGTGCGACGAGAGAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
DB 518 GGTGCGACGAGAGAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 459
QY 361 GAGACAGCTGTGTGAAGAACATATGTCGTGTGACGCGCTGTGCGCAAGAGAGACT 419
DB 488 GAGACAGCTGTGTGAAGAACATATGTCGTGTGACGCGCTGTGCGCAAGAGAGACT 399
QY 420 AAGGAGAGGAGAGACTATGTGTGAGCTTTTAAATAGAGAGATGACTCGAATTGAGT 479
DB 398 AAGGAGAGGAGAGACTATGTGTGAGCTTTTAAATAGAGAGATGACTCGAATTGAGT 339
QY 480 GATCATTTAGAGCTGAGAGCTCTGTTCTCTGAGAGAGTGAAGAGAGAGAGAGAG 539
DB 338 GATCATTTAGAGCTGAGAGCTCTGTTCTCTGAGAGAGTGAAGAGAGAGAGAGAG 279
QY 540 GGAATGAGGTTTGTGAGAAATCTCTGAGAGAGCTCTCTGCAATGAGCTGAGATGAG 599
DB 278 GGAATGAGGTTTGTGAGAAATCTCTGAGAGAGCTCTCTGCAATGAGCTGAGATGAG 219
QY 600 CAGACGCCCGAGATTGTTCTCTGAGAGAGATTTCTTTCACAGATGAGAGATTCTTT 639
DB 218 CAGACGCCCGAGATTGTTCTCTGAGAGAGATTTCTTTCACAGATGAGAGATTCTTT 159
QY 660 GCTTATGTTGAATTCATGCTCTCTTCTCATACAGAGAGATGAGATGAGATGCTTCT 719
DB 158 GCTTATGTTGAATTCATGCTCTCTTCTCATACAGAGAGATGAGATGAGATGCTTCT 99
QY 720 TTGTTGTCTGATTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTT 779
DB 98 TTGTTGTCTGATTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTT 39
QY 780 AAAGAGAGAAAGTAAATGTACAGTTTAAATGAG 817
DB 38 AAAGAGAGAAAGTAAATGTACAGTTTAAATGAG 1

```

RESULT 2

AL572703/c
LOCUS
DEFINITION
AL572703 LTI_NFL006 Pl2 Homo sapiens cDNA clone CS0D1035YE24 3
prime, mRNA sequence.
ACCESSION
AL572703
VERSION
AL572703.1 GI:12911228
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 910)
AUTHORS
Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished (2001)
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1..910
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1035YE24"
/clone_id="LTI_NFL006 Pl2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com>"

BASE COUNT 244 a 251 c 214 g 195 t 6 others
ORIGIN

Query Match 92.0%; Score 802.8; DB 9; Length 910;
Best Local Similarity 99.3%; Pred. No. 1.4e-136;
Matches 812; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

```

QY 1 CTCACGATATGTTCAACTATGAAGAATCTGACCGCCCAACGCACTGCTGCGCTTG 60
DB 824 CTCACGATATGTTCAACTATGAAGAATCTGACCGCCCAACGCACTGCTGCGCTTG 765
QY 61 CCGTGATCTCTCCACGCTGCTGCTTGTACGTGAGAGAGAACTCTGCAATTAATTCT 120
DB 764 CCGTGATCTCTCCACGCTGCTGCTTGTACGTGAGAGAGAACTCTGCAATTAATTCT 705
QY 121 CTATGAGAGCTGCGGGGCAATAAGAACACTACCGCTCTGAGAGAGCTGCAATGCTCCG 180
DB 704 CTATGAGAGCTGCGGGGCAATAAGAACACTACCGCTCTGAGAGAGCTGCAATGCTCCG 645
QY 181 CTGCTTCCGACAGAGAGAAATCCCTCCCTGCTTGCTGCTCAAGGTGTGTCTGCG 240
DB 644 CTGCTTCCGACAGAGAGAAATCCCTCCCTGCTTGCTGCTCAAGGTGTGTCTGCG 585
QY 241 GGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
DB 584 GGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 525
QY 301 GGTGCGACGAGAGAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
DB 524 GGTGCGACGAGAGAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 465
QY 361 GAGACAGCTGTGTGAAGAACATATGTCGTGTGACGCGCTGTGCGCAAGAGAGACT 419
DB 464 GAGACAGCTGTGTGAAGAACATATGTCGTGTGACGCGCTGTGCGCAAGAGAGACT 405
QY 420 AAGGAGAGGAGAGACTATGTGTGAGCTTTTAAATGAGAGAGATGACTCGAATTGAGT 479

```

Db 404 AAGGAGGAGGAGACTATGTGTAGCTTTTAAATAGAGGATTGACTCGGATTTGAGT 345

Qy 480 GATCATTAGGAGCTGAGGCTGTTTCTCGGAGGTAGAGACGGCTGTTCTGCTGCA 539

Db 344 GATCATTAGGAGCTGAGGCTGTTTCTCGGAGGTAGAGACGGCTGTTCTGCTGCA 285

Qy 540 GGGATGGGTTTCTTGGAAATCTCTAGAGAGCTCTCTCTGCGATGCTGCACTCG 599

Db 284 GGGATGGGTTTCTTGGAAATCTCTAGAGAGCTCTCTCTGCGATGCTGCACTCG 225

Qy 600 CAGAGAGCCCGAGTGTCTCTGCGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 659

Db 224 CAGAGAGCCCGAGTGTCTCTGCGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 165

Qy 660 GCTTATGTGAATTCATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 719

Db 164 GCTTATGTGAATTCATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 105

Qy 720 TTTGTTGTCTGATTTATGTTTATTAAGTAAACAAAGTTTATTAAGATTTCTG 779

Db 104 TTTGTTGTCTGATTTATGTTTATTAAGTAAACAAAGTTTATTAAGATTTCTG 45

Qy 780 AAAGAGGAAAGTAAATGTACAACTTAAATAAAG 817

Db 44 AAAGAGGAAAGTAAATGTACAACTTAAATAAAG 7

RESULT 3

AL548687/c 1001 bp mRNA linear EST 16-FEB-2001

LOCUS AL548687.LPI.NFL006.PL2 Homo sapiens cDNA clone GSOD1036P13 3

DEFINITION prime, mRNA sequence.

ACCESSION AL548687

VERSION AL548687.1 GI:12883941

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Li, W.B., Gruber, C., Jesse, J. and Polayes, D.

JOURNAL Full-length cDNA libraries and normalization

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage

Bp 191 91006 EVRY cedex - France

Email: segre@genoscope.cns.fr, Web: www.genoscope.cns.fr.

FEATURES

source

1..1001

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_id="LPI.NFL006.PL2"

/clone_lib="LPI.NFL006.PL2"

/issue_type="placenta"

/note="Vector: pCMVSORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact: Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax: (1) 301 610 8371

Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com"

BASE COUNT 265 a 279 c 237 g 217 t 3 others

ORIGIN

Query Match 89.9%; Score 785; DB 9; Length 1001;

Best Local Similarity 99.8%; Pred.No.2.2e-133;

Matches 807; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Qy 1 CTCAGGAGATGTTCACATGAGATATGTCACCGGACGAGTCACATGGGCTTG 60

Db 809 CTCAGGAGATGTTCACATGAGATATGTCACCGGACGAGTCACATGGGCTTG 750

Qy 61 CCGTGACATCCCTCCACGCTGCTGATCTTTGACGTGAGAGAACTCCCAATTAATTCAT 120

Db 749 CCGTGACATCCCTCCACGCTGCTGATCTTTGACGTGAGAGAACTCCCAATTAATTCAT 630

Qy 121 CTATGAGAGCTGCCGGGGCAATTAAGAACAGTACCGCTCTGAGAGAGCTCTGATCCG 180

Db 689 CTATGAGAGCTGCCGGGGCAATTAAGAACAGTACCGCTCTGAGAGAGCTCTGAT 631

Qy 181 CTGCTCCGACAGAGAAATCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 240

Db 630 CTGCTCCGACAGAGAAATCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 571

Qy 241 GGGGCTGTGTATGATGATGATCTCTCTCTGAGAGAGCTCCAGAGTCTGATCCG 300

Db 570 GGGGCTGTGTATGATGATGATCTCTCTCTGAGAGAGCTCCAGAGTCTGATCCG 511

Qy 301 GGTGACAGAGAGAACAGAGAGCTGCCCTGCGACCCGCTGAGCTCCGAGATGACAA 360

Db 510 GGTGACAGAGAGAACAGAGAGCTGCCCTGCGACCCGCTGAGCTCCGAGATGACAA 451

Qy 361 GAGAGAGCTGTGTAAGAACACATATGTCCTGTGACCGGCTGTGCGAAGAGACT 636 419

Db 450 GAGAGAGCTGTGTAAGAACACATATGTCCTGTGACCGGCTGTGCGAAGAGACT 391

Qy 420 AAGGAGGAGGAGACTATGTGAGCTTTTAAATAGAGGATGATGATCGGATTTGAGT 479

Db 390 AAGGAGGAGGAGACTATGTGAGCTTTTAAATAGAGGATGATGATCGGATTTGAGT 331

Qy 480 GATCATTAGGAGCTGAGAGCTGTTTCTCGGAGGTAGAGACGGCTGTTCTGCTGCA 539

Db 330 GATCATTAGGAGCTGAGAGCTGTTTCTCGGAGGTAGAGACGGCTGTTCTGCTGCA 271

Qy 540 GGGATGGGTTTCTTGGAAATCTCTAGAGAGCTCTCTCTGCAATGCTGCAATCTGG 599

Db 270 GGGATGGGTTTCTTGGAAATCTCTAGAGAGCTCTCTCTGCAATGCTGCAATCTGG 211

Qy 600 CAGAGAGCCCGAGTGTCTCTGCGATGCTATTTCTTCTCTCTCTCTCTCTCTCTCT 659

Db 210 CAGAGAGCCCGAGTGTCTCTGCGATGCTATTTCTTCTCTCTCTCTCTCTCTCTCT 151

Qy 660 GCTTATGTGAATTCATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 719

Db 150 GCTTATGTGAATTCATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 91

Qy 720 TTTGTTGTCTGATTTATGTTTATTAAGTAAACAAAGTTTATTAAGATTTCTG 779

Db 90 TTTGTTGTCTGATTTATGTTTATTAAGTAAACAAAGTTTATTAAGATTTCTG 31

Qy 780 AAAGAGGAAAGTAAATGTACAACTTAA 808

Db 30 AAAGAGGAAAGTAAATGTACAACTTAA 2

RESULT 4

BE621676/c 923 bp mRNA linear EST 20-OCT-2000

LOCUS BE621676 60149368371 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:3895701 3'

DEFINITION mRNA sequence.

ACCESSION BE621676

VERSION BE621676.1 GI:9892714

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 923)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov

Plate: LLM9687 row: f column: 22

High quality sequence start: 12

High quality sequence stop: 771.

FEATURES

Location/Qualifiers

1..923

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:3895701"

/clone_lib="NIH MGC 70"

/tissue_type="epithelioid carcinoma"

/lab_host="DH10B (phage-resistant)"

/note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr.
 Average insert size 1.1 kb. Library constructed by Life
 Technologies."

BASE COUNT

244 a 258 c 220 g 201 t

ORIGIN

Query Match 89.7%; Score 783.4; DB 10; Length 923;

Best Local Similarity 98.8%; Pred. No. 4,5e-133;

Matches 831; Conservative 0; Mismatches 6; Indels 4; Gaps 4;

QY 1 CTCGAGGATATGTTCAACTATGAAAGATCTGACCCGCAACGAGTCACTGGGCTT 59
 DB 843 CTCGAGGAAATGTCAACTATGAAAGAA-ACCTGACCGCAACGAGTCACTGGGCTT 785
 QY 60 GCGGTGATCTCTCCACGCTGCTTGAAGTGAAGAGAACTCTGAGAGAGGCTGATCTCC 119
 DB 784 GCGGTGATCTCTCCACGCTGCTTGAAGTGAAGAGAACTCTGAGAGAGGCTGATCTCC 725
 QY 120 TCTATGAGAGCTGCGGGGCAATAGAAAGAGTACCGCTCTGAGAGAGGCTGATCTCC 179
 DB 724 TCTATGAGAGCTGCGGGGCAATAGAAAGAGTACCGCTCTGAGAGAGGCTGATCTCC 665
 QY 180 GCTGCTTCGCGCAAGAGAGAAATCTCCCTGCGCTTGGCTCAAAAGTGGTCTTCC 239
 DB 664 GCTGCTTCGCGCAAGAGAGAAATCTCCCTGCGCTTGGCTCAAAAGTGGTCTTCC 605
 QY 240 CGGGGCTGTTGATGATGATGATCTCTCTGAGAGAGGCTGATCTGATCTCC 299
 DB 604 CGGGGCTGTTGATGATGATGATCTCTCTGAGAGAGGCTGATCTGATCTCC 545
 QY 300 GGGTGGCAGAGAGAAACAGAGAGGCTGCGACCGCTGAGAGCTCCGAGATGACA 359
 DB 544 GGGTGGCAGAGAGAAACAGAGAGGCTGCGACCGCTGAGAGCTCCGAGATGACA 485
 QY 360 AGAGAGAGCTGTGAAGAAACATATGCTCTGACCGCTCTGCGCAAGAGAGCT 418
 DB 484 AGAGAGAGCTGTGAAGAAACATATGCTCTGACCGCTCTGCGCAAGAGAGCT 425
 QY 419 GAAGAGAGGAGAGACTATGTGAGAGCTTTTAAATAGAGAGATGATCGATTTGAG 478
 DB 424 GAAGAGAGGAGAGACTATGTGAGAGCTTTTAAATAGAGAGATGATCGATTTGAG 365
 QY 479 TGATCATTTAGGGCTAGGCTGTTTCTCTGGAGAGTGAAGAGCGCTGCTTCTGGCT 538
 DB 364 TGATCATTTAGGGCTAGGCTGTTTCTCTGGAGAGTGAAGAGCGCTGCTTCTGGCT 305
 QY 539 AGGATAGGGCTGTTTGAAGAAATCTCTAGAGAGCTCTCTCTGAGAGCTGAGCTG 598
 DB 304 AGGATAGGGCTGTTTGAAGAAATCTCTAGAGAGCTCTCTCTGAGAGCTGAGCTG 245
 QY 599 GAGAGAGCCCGAGTGTCTCTGAGAGAGATTTCTTCTCCAGGATGAGATTTCTT 658
 DB 244 GAGAGAGCCCGAGTGTCTCTGAGAGAGATTTCTTCTCCAGGATGAGATTTCTT 185

QY 659 TGCTATATTTGAATTCATTCGCTCTTTCTCATCATAGAGAGTATGTGGATCGTTTC 718
 DB 184 TGCTATATTTGAATTCATTCGCTCTTTCTCATCATAGAGAGTATGTGGATCGTTTC 125
 QY 719 TTTTGTCTTCGATTTATGTTTATTTTAAATGATTAACAAAGTTTATTTGATCTTC 778
 DB 124 TTTTGTCTTCGATTTATGTTTATTTTAAATGATTAACAAAGTTTATTTGATCTTC 65
 QY 779 GAAAGAGAGAAAGTAAATGTACAGATTATTAATAAAAGGAGGCTTCCCTTGAATATA 838
 DB 64 GAAAGAGAGAAAGTAAATGTACAGATTATTAATAAA-TACGCTTCCCTTGAATATA 6
 QY 839 T 839
 DB 5 T 5

RESULT 5
 AL575665/c AL575665 1061 bp mRNA linear EST 16-FEB-2001
 LOCUS AL575665 LTI_NFL006_PU2 Homo sapiens cDNA clone CSOD1069Y17 3
 DEFINITION prime, mRNA sequence.
 ACCESSION AL575665
 VERSION AL575665.1 GI:12937050
 KEYWORDS EST.
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 1061)
 AUTHORS Li'W.B., Gruber,C., Jesse,J. and Polayes,D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 FEATURES
 source
 1..1061
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CSOD1069Y17"
 /clone_lib="LTI_NFL006_PU2"
 /tissue_type="placenta"
 /note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-stranded cDNA was digested with Not I and
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6
 vector. Library was normalized. Library was constructed by
 Life Technologies. Contact : Feng Liang Life Technologies,
 a division of Invitrogen 9600 Medical Center Drive
 Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
 Email: fliang@lifetech.com URL :
 http://fulllength.invitrogen.com"

BASE COUNT 281 a 301 c 250 g 226 t 3 others
 ORIGIN

Query Match 89.6%; Score 782.4; DB 9; Length 1061;
 Best Local Similarity 99.5%; Pred. No. 6,4e-133;
 Matches 805; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 1 CTCGAGGATATGTTCAACTATGAAAGATCTGACCCGCAACGAGTCACTGGGCTTG 60
 DB 808 CTCGAGGATATGTTCAACTATGAAAGATCTGACCCGCAACGAGTCACTGGGCTTG 749
 QY 61 CCGTGATTCCTTCCACGCTGTGATCTTGAAGAGAACTCTGCAATACTTCAT 120
 DB 748 CCGTGATTCCTTCCACGCTGTGATCTTGAAGAGAACTCTGCAATACTTCAT 689
 QY 121 CTATGAGAGCTGCGGGGCAATAGAAAGAGCTACGCTCTGAGAGAGGCTGATCTCCG 180
 DB 688 CTATGAGAGCTGCGGGGCAATAGAAAGAG-TACGCTCTGAGAGAGGCTGATCTCCG 630

```

Qy 181 CTGCTCCGCCACAGAGAAATCCCTCCCTGAGCTCAAGTGTGTCTGAC 240
Db 629 CTGCTCCGCCACAGAGAAATCCCTCCCTGAGCTCAAGTGTGTCTGAC 570
Qy 241 GGGGCTGTGTGATGCTGTATCTCTTCTTCTGAGAGCTCCATGCTCACTGATCCG 300
Db 569 GGGGCTGTGTGATGCTGTATCTCTTCTTCTGAGAGCTCCATGCTCACTGATCCG 510
Qy 301 GGTGACCGGAGAACACAGAGAGTGCCTGCGACCGCTGTGAGTCCGAGATGACAA 360
Db 509 GGTGACCGGAGAACACAGAGAGTGCCTGCGACCGCTGTGAGTCCGAGATGACAA 450
Qy 361 GAGAGAGCTGTGAGAGAACATATGTCCTGTACCCCTGTGCGCAAGAGACT- GGG 419
Db 449 GAGAGAGCTGTGAGAGAACATATGTCCTGTACCCCTGTGCGCAAGAGACTGGGG 330
Qy 420 AAGGAGGAGAGACTATGTGTAGCTTTTAAATAGAGGATTGACTCGGATTTGAGT 479
Db 389 AAGGAGGAGAGACTATGTGTAGCTTTTAAATAGAGGATTGACTCGGATTTGAGT 330
Qy 480 GATCATTAAGGCTGAGAGTCTGTTCTCTGAGAGTGAAGAGGCTGTTCTGTGCA 539
Db 329 GATCATTAAGGCTGAGAGTCTGTTCTCTGAGAGTGAAGAGGCTGTTCTGTGCA 270
Qy 540 GGAATGAGTGTGCTTTGAGAAATCCCTCTAGAGAGCTCCTGCGAATGGCTGAGTCCG 599
Db 269 GGAATGAGTGTGCTTTGAGAAATCCCTCTAGAGAGCTCCTGCGAATGGCTGAGTCCG 210
Qy 600 CAGAGAGCCCGAGTGTCTCTGCTGATGATTTCTTCTCTCAAGTATGAGTTCCTT 659
Db 209 CAGAGAGCCCGAGTGTCTCTGCTGATGATTTCTTCTCTCAAGTATGAGTTCCTT 150
Qy 660 GCTATGTGAATTCATTCCTCTTCTCATACAGAGTGAATGATGTTCT 719
Db 149 GCTATGTGAATTCATTCCTCTTCTCATACAGAGTGAATGATGTTCT 90
Qy 720 TTGTTGTCTGATTTATGTTTAAATAGATTAAGATTAAGATTCG 779
Db 89 TTGTTGTCTGATTTATGTTTAAATAGATTAAGATTAAGATTCG 30
Qy 780 AAGAGAGAAATTAATGTACAGTTTA 808
Db 29 AAGAGAGAAATTAATGTACAGTTTA 1

RESULT 6
AUS73986/c 983 bp mRNA linear EST 16-FEB-2001
LOCUS AUS73986
DEFINITION AUS73986 LTR_NFL006_PL2 Homo sapiens cDNA clone CS0D1053YK07 3
prime, mRNA sequence.
ACCESSION AUS73986
VERSION AUS73986.1 GI:12933753
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 983)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
Source
1..983
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1053YK07"
/clone.lib="LTR_NFL006_PL2"
/risue.type="placenta"
/Note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA

```

was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetechn.com URL : <http://fulllength.invitrogen.com>

BASE COUNT 258 a 276 c 230 g 212 t 7 others
ORIGIN

```

Query Match 89.5% Score 781.6; DB 9; Length 983;
Best Local Similarity 99.1%; Pred. No. 9.3e-133;
Matches 803; Conservative 3; Mismatches 2; Indels 2; Gaps 2;

Qy 1 CTCAGGATATATGTTCACTATGAGAAATACATGACCCGCAACGACAGTCACTGAGCCCTTG 60
Db 811 CTCAGGATATATGTTCACTATGAGAAATACATGACCCGCAACGACAGTCACTGAGCCCTTG 752
Qy 61 CCGTGCATCCTTCCACGCTGATCTTTGACGTGAGAGAACTCCTGCATTAATTCTAT 120
Db 751 CCGTGCATCCTTCCACGCTGATCTTTGACGTGAGAGAACTCCTGCATTAATTCTAT 692
Qy 121 CTATGAGAGGCTGCGGGGCAATPAAGAA-CAGCTAACGCTCTGAGAGAGGCTGCAATGCTCC 179
Db 691 CTATGAGAGGCTGCGGGGCAATPAAGAACTGAGAGAGGCTGCAATGCTCC 632
Qy 180 GCTGCTTCCGACAGAGAAATCCTCCCTGAGCCCTTGGCTCAAGAGGAGTGTCTG 239
Db 631 GCTGCTTCCGACAGAGAAATCCTCCCTGAGCCCTTGGCTCAAGAGGAGTGTCTG 572
Qy 240 CGGGAGCTGTGTGATGTGTGATCCTTCTCTGAGAGCTTCATAGTCTTACTGATCC 299
Db 571 CGGGAGCTGTGTGATGTGTGATCCTTCTCTGAGAGCTTCATAGTCTTACTGATCC 512
Qy 300 GGGTGGCAAGAGAGAAACAGAGAGCGTGCCTGCGCAACCGTCTGAGAGCTCCGAGATGACA 359
Db 511 GGGTGGCAAGAGAGAAACAGAGAGCGTGCCTGCGCAACCGTCTGAGAGCTCCGAGATGACA 452
Qy 360 AGGAGAGCTGTGTGAGAGACATATGTCCTGTGACCGCCCTGTGCGCAAGAGACT- GG 418
Db 451 AGGAGAGCTGTGTGAGAGACATATGTCCTGTGACCGCCCTGTGCGCAAGAGACTGGG 392
Qy 419 GAAGGAGGGGAGACTATGTGTGAGCTTTTAAATAGAGGATGACTCGGATTTGAG 478
Db 391 GAAGGAGGGGAGACTATGTGTGAGCTTTTAAATAGAGGATGACTCGGATTTGAG 332
Qy 479 TGATCATTAAGGCTGAGGCTGTCTTCTCTGAGAGATGAGAGCGTCTCTCTGTGCGC 538
Db 331 TGATCATTAAGGCTGAGGCTGTCTTCTCTGAGAGATGAGAGCGTCTCTCTGTGCGC 272
Qy 539 AGGAGTGGGTTTCTTGTGAAATCCTCTAGAGAGCTCCTCTCGCATGAGCCGCAAGCTG 598
Db 271 AGGAGTGGGTTTCTTGTGAAATCCTCTAGAGAGCTCCTCTCGCATGAGCCGCAAGCTG 212
Qy 599 GCAGAGCCCGAGTGTGTTCTCTGCTGATGATTTCTTCTCTGAGAGGTTTCTT 658
Db 211 GCRGAGCCCGAGTGTGTTCTCTGCTGATGATTTCTTCTCTGAGAGGTTTCTT 152
Qy 659 TGCTTATGTTGAATTCATTCCTCTTCTCTCATACAGAGTGAATGTGAGTCTTC 718
Db 151 KGCTTATGTTGAATTCATTCCTCTTCTCTCATACAGAGTGAATGTGAGTCTTC 92
Qy 719 TTGTTGTGCTATTTATGTTTAAATAGATTAAGATTAAGATTCG 778
Db 91 TTGTTGTGCTATTTATGTTTAAATAGATTAAGATTAAGATTCG 32
Qy 779 GAAGAGAGAAATTAATGTACAGTTTA 808
Db 31 GAAGAGAGAAATTAATGTACAGTTTA 2

```

```

RESULT 7
AL575631/c 938 bp mRNA linear EST 16-FEB-2001
LOCUS AL575631 LTI NF1006 Pl2 Homo sapiens cDNA clone CS0D1069Y006 3
DEFINITION prime, mRNA sequence.
ACCESSION AL575631
VERSION AL575631.1 GI:12936982
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 938)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006-EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..938
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1069Y006"
/clone_1b="LTI NF1006 Pl2"
/tissue_type="placenta"
/notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 250 a 260 c 223 g 201 t 4 others
ORIGIN
Query Match 88.9%; Score 776.2; DB 9; Length 938;
Best Local Similarity 99.5%; Pred. No. 9,2e-132;
Matches 797; Conservative 2; Mismatches 0; Indels 2; Gaps 2;

QY 2 TCACGAGATATGTTCACTATGAAGAACTACTGACCGCAACGAGTCACTGGGCTTGC 61
DB 800 TCACGAGATATGTTCACTATGAAGAACTACTGACCGCAACGAGTCACTGGGCTTGC 741
QY 62 CGTGCACTCTCCCAAGCTGGTACTTGAAGTGAAGAGAACTCTGCAATTAATTGATC 121
DB 740 CGTGCACTCTCCCAAGCTGGTACTTGAAGTGAAGAGAACTCTGCAATTAATTGATC 681
QY 122 TATGAGAGCTGCGCGGCAATTAAGAGAGTACCTCTGAGAGAGCTTGATGCTCCG 181
DB 680 TATGAGAGCTGCGCGGCAATTAAGAGAGTACCTCTGAGAGAGCTTGATGCTCCG 622
QY 182 TGGTCCGCGCAGAGAGATCTCCCTGCGCTGAGAGTGAAGTGGTGGTGGGCG 241
DB 621 TGGTCCGCGCAGAGAGATCTCCCTGCGCTGAGAGTGAAGTGGTGGTGGGCG 562
QY 242 GGGCTGTTGATGATGTTGATCTCTTCTGAGAGCTCCATGATGTTACTGATCCG 301
DB 561 GGGCTGTTGATGATGTTGATCTCTTCTGAGAGCTCCATGATGTTACTGATCCG 502
QY 302 GTGGCAGCGAGAGAACCAAGAGCGTCCCTGCGACCTGTTGAGAGTCCGAGATGCAAG 361
DB 501 GTGGCAGCGAGAGAACCAAGAGCGTCCCTGCGACCTGTTGAGAGTCCGAGATGCAAG 442
QY 362 GAGCAGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
DB 441 GAGCAGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 382
QY 421 AGGAGAGGAGACTATGTTGAGCTTTTAAATGAGAGATTGACTCGGATTTGATG 480

```

```

DB 381 AGGAGAGGAGACTATGTTGAGCTTTTAAATGAGAGATTGACTCGGATTTGATG 322
QY 481 ATCATTAAGAGGCTGAGTGTCTGTTCTCTGAGAGTGAAGACGAGTCTCTGCTGAGCG 540
DB 321 ATCATTAAGAGGCTGAGTGTCTGTTCTCTGAGAGTGAAGACGAGTCTCTGCTGAGCG 262
QY 541 GGATGGGTTTGTCTTGAATCTCTGAGAGGCTCTCTGATGAGTGGCTGAGTCTGAGC 600
DB 261 GGATGGGTTTGTCTTGAATCTCTGAGAGGCTCTCTGATGAGTGGCTGAGTCTGAGC 202
QY 601 AGAGAGCCGAGTGTCTCTGCGCTGATGATTTCTTCCCGAGGATGATGTTCTTTG 660
DB 201 AGAGAGCCGAGTGTCTCTGCGCTGATGATTTCTTCCCGAGGATGATGTTCTTTG 142
QY 661 CTAAATGGAATTCATTCCTCTTCTTCATCAGAGAGTGAATGATGTTCTT 720
DB 141 CTAAATGGAATTCATTCCTCTTCTTCATCAGAGAGTGAATGATGTTCTT 82
QY 721 TTGTTGTCTGATTTATGTTTATTTTAAATTAACATTAACAAAGTTTATTAAGATTTGTA 780
DB 81 TTGTTGTCTGATTTATGTTTATTTTAAATTAACATTAACAAAGTTTATTAAGATTTGTA 22
QY 781 AAGAGAGAAAGTAAATGTAC 801
DB 21 AAGAGAGAAAGTAAATGTAC 1

RESULT 8
AL573346/c 1052 bp mRNA linear EST 16-FEB-2001
LOCUS AL573346 LTI NF1006 Pl2 Homo sapiens cDNA clone CS0D1042YC20 3
DEFINITION prime, mRNA sequence.
ACCESSION AL573346
VERSION AL573346.1 GI:12932499
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1052)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006-EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..1052
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1042YC20"
/clone_1b="LTI NF1006 Pl2"
/tissue_type="placenta"
/notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 279 a 298 c 248 g 222 t 5 others
ORIGIN
Query Match 88.9%; Score 776; DB 9; Length 1052;
Best Local Similarity 99.5%; Pred. No. 9,4e-132;
Matches 798; Conservative 1; Mismatches 1; Indels 2; Gaps 2;

QY 1 CTCACGATATGTTCACTATGAAGAACTGACCGCAACGAGTCACTGGGCTTGG 60

```

Db	802	CTCCACGCAATATGTTCAACTTGAAGAAATCTGCAACGCCACAGCAGTCACTGGGCTTGG	743
Qy	61	CCGTGATACCTTCCACACGCTGTACTTTATACGTGAGAGAGAACTCTCGAATAATCTCAT	120
Db	742	CCGTGATACCTTCCACACGCTGTACTTTATACGTGAGAGAGAACTCTCGAATAATCTCAT	683
Qy	121	CTATGAGAGCTCCCGGGCAATTAAG-AAAGCTTACCGCTCTGAGAGAGGCTGCATGCTCC	179
Db	682	CTATGAGAGCTCCCGGGCAATTAAGAAACAGCTTACCGCTCTGAGAGAGGCTGCATTTTCC	623
Qy	180	GCAGCTTCCGCGCAGAGAGAAATCCCTCCCTCCCTTGGGTCAAAAGTGTGTCTTCG	239
Db	622	GCAGCTTCCGCGCAGAGAGAAATCCCTCCCTCCCTTGGGTCAAAAGTGTGTCTTCG	563
Qy	240	CGGAGCTGTTCCTGATGCTGTATCTCTTCTGGAGGCTCCAGTGTCTACGATCC	299
Db	562	CGGAGCTGTTCCTGATGCTGTATCTCTTCTGGAGGCTCCAGTGTCTACGATCC	503
Qy	300	GGGTGCACGAGAGAACCCAGAGCGTCCCTGCGACCGCTGTGAGCTCCGAGATGACA	359
Db	502	GGTGTGCACGAGAGAACCCAGAGCGTCCCTGCGACCGCTGTGAGCTCCGAGATGACA	443
Qy	360	AGAGAGAGCTGTGTAAGAAACATATGCTCTGTACCGGCTCTGCGCAAGAGACT-GG	418
Db	442	AGAGAGAGCTGTGTAAGAAACATATGCTGTGACCGGCTCTGCGCAAGAGACTGGG	383
Qy	419	GAAGGAGAGGAGACATATGCTGAGCTTTTTTAATAGAAGGATTGATCGGATTTGAG	478
Db	382	GAAGGAGAGGAGACATATGCTGAGCTTTTTTAATAGAAGGATTGATCGGATTTGAG	323
Qy	479	TGATCTTAGAGGCTGAGGCTGTGTTCTCGGAGGATGAGACGGCTCTTCTGTCTGAGC	538
Db	322	TGATCTTAGAGGCTGAGGCTGTGTTCTCGGAGGATGAGACGGCTCTTCTGTCTGAGC	263
Qy	539	AGGGAATGGGTTTGTCTTGGAAATCTCTTGAGAGGCTCTCTCTGCAATGCTCGAGTCTG	598
Db	262	AGGGAATGGGTTTGTCTTGGAAATCTCTTGAGAGGCTCTCTCTGCAATGCTCGAGTCTG	203
Qy	599	GCAGCAGCCCGAGTGTCTCTGCTGATGCAATTTCTTCCGACGATGAGTTTCTT	658
Db	202	GCAGCAGCCCGAGTGTCTCTCTGCTGATGCAATTTCTTCCGACGATGAGTTTCTT	143
Qy	659	TGCTTATGTTGATTTCCATTGCTCTCTTCTGATCAGAGAGATGTTGATGCTTTC	718
Db	142	TGCTTATGTTGATTTCCATTGCTCTCTTCTGATCAGAGAGATGTTGATGCTTTC	83
Qy	719	TTTTGTTTGTCTGATTTATGTTTTTTAAGTATMAACAAAGTTTTTTATAGCATCT	778
Db	82	TTTTGTTTGTCTGATTTATGTTTTTTAAGTATMAACAAAGTTTTTTATAGCATCT	23
Qy	779	GAAGAGAGAGAGTAAATGTA 800	
Db	22	GAAGAGAGAGAGTAAATGTA 1	

Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
 source
 1. 877
 /organism="Homo sapiens"
 /db_xref="taxon:3608"
 /clone="CS0D1075YC08"
 /clone_1b="UT1_NFL006_PL12"
 /tissue_type="Placenta"
 /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : Elliang@lifetechn.com URL : <http://fulllength.invitrogen.com>"

BASE COUNT 226 a 245 c 205 g 192 t 9 others

Query Match 88.3%; Score 770.8; DB 9; Length 877;
 Best Local Similarity 98.1%; Pred. No. 9..1e-131;
 Matches 783; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

QY 1 CTCGAGCAATGTTCAACTATGAAAGAACTATGACCCGCCAAGCAAGCACTGGGCGCTTG 60
 Db 809 CTCACGCAATATGTTCACTATGAAAGAACTATGACCCGCCAAGCAAGCACTGGGCGCTTG 750
 QY 61 CCGTGCACTCTTCCCAAGCGTGTACTTTGACGTGAGAGAGAACTCTTGCAATACTTCAT 120
 Db 749 CCGTGCACTCTTCCCAAGCGTGTACTTTGACGTGAGAGAGAACTCACAAATACTTCAT 690
 QY 121 CTATGAGAGGCTGCCG396CAATPAAGACAAGTACCGCTCTGAGAGAGGCGTGCATGCTCG 180
 Db 689 CTATGAGAGGCTGCCG396CAATPAAGACAAGTACCGCTCTGAGAGAGGCGTGCATGCTCG 630
 QY 181 CAGTCTCCGACAGAGAGAAATCTCCCTCCCTGCGCTTGGTCAAAAGTGGTGGTTCGCG 240
 Db 629 CAGTCTCCGACAGAGAGAAATCTCCCTCCCTGCGCTTGGTCAAAAGTGGTGGTTCGCG 570
 QY 241 GGGGCTGTTCTGATGAGTGTATCTCTTCTGAGAGCGTCCATAGGTCTACCTGCATCCG 300
 Db 569 GGGGCTGTTCTGATGAGTGTATCTCTTCTGAGAGCGTCCATAGGTCTACCTGCATCCG 510
 QY 301 GGTGGCAAGAGAAACAGAGAGCGTCCCTTCGACACGCTCTGAGAGCTCCGAGATGACAA 360
 Db 509 GGTGGCAAGAGAAACAGAGAGCGTCCCTTCGACACGCTCTGAGAGCTCCGAGATGACAA 450
 QY 361 GAGAGAGCTGTGTAAGAACACATATGTCTGTGACGCGCTGTCCCAAGAGACT- GGG 419
 Db 449 GAGAGAGCTGTGTAAGAACACATATGTCTGTGACGCGCTGTCCCAAGAGACTGGGG 390
 QY 420 AAGGAGAGGAGAGACTATGTGTAGAGCTTTTAAATAGAGGGATTGATCGGATTTGAGT 479
 Db 389 AAGGAGAGGAGAGACTATGTGTAGAGCTTTTAAATAGAGGGATTGATCGGATTTGAGT 330
 QY 480 GATCATTTAGGAGCTGAGGCTGTGTTTCTCTGAGAGAGTGAAGACGCTGCTTCTGTGCTGCA 539
 Db 329 GATCATTTAGGAGCTGAGGCTGTGTTTCTCTGAGAGAGTGAAGACGCTGCTTCTGTGCTGCA 270
 QY 540 GGGATGGGTTGCTTTGAAATCTCTTAGAGAGCTCTCTCTGATAGGAGCTGACGTGG 599
 Db 269 GGGATGGGTTGCTTTGAAATCTCTTAGAGAGCTCCACTCGATGGAGCTGACGTGG 210
 QY 600 CAGCAGCCCCAGATGTTTCTCTGCGCTGATGATTTCTTTTCTCCAGTAGAGTTTCTTT 659
 Db 209 CAGCAGCCCCAGATGTTTCTCTGCGCTGATGATTTCTTTTCTCCAGTAGAGTTTCTTT 150
 QY 660 GCTTATGTTGATTCATGCTCTTTTCTCATACAGAAAGTAGAGTTGAGATGCTTTCT 719

Db 149 GCTTAGTGTGAATTCATCGATCTTTCTCATCAGAGATGATGTGGATCGTGTCT 90

Qy 720 TTTGTTGTCGATTTATGATTTTATTAAGTATTAACAAAAGCTTTTATTAGATTCG 779

Db 89 TTTGTTGTCGATTTATGATTTTATTAAGTATTAACAAAAGCTTTTATTAGATTCG 30

Qy 780 AAAGAGAGAAAGTAAAT 797

Db 29 AAAGAGAGAAAGTAAAT 12

RESULT 10
AL576831/c 950 bp mRNA linear EST 16-FEB-2001
LOCUS AL576831 LTI_NFL006_P12 Homo sapiens cDNA clone CS0D1079B07 3
DEFINITION prime, mRNA sequence.
ACCESSION AL576831
VERSION AL576831.1 GI:12939362
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 950)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1..950
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1079B07"
/clone_lib="LTI_NFL006_P12"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email: fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 252 a 263 c 224 g 208 t 3 others

ORIGIN

Query Match 88.3%; Score 770.6; DB 9; Length 950;
Best Local Similarity 99.3%; Pred. No. 9,5e-131;
Matches 803; Conservative 2; Mismatches 1; Indels 3; Gaps 3;

Qy 1 CTCAGGATATGTTCAACTAAGAGATCTGACCGGCAACGAGTCACTGAGCCTTG 60

Db 808 CTCAGGATATGTTCAACTAAGAGATCTGACCGGCAACGAGTCACTGAGCCTTG 749

Qy 61 CCGTGATCTCTCCACGCTGTACTTTGACGTGAGAGAGAACTCTGCAATTAATTCA 120

Db 748 CCGTGATCTCTCCACGCTGTACTTTGACGTGAGAGAGAACTCTGCAATTAATTCA 689

Qy 121 CTATGAGAGCTGCGGGGCAATTAAGAACACTACCGCTGAGAGGCTGCAATGCTCCG 180

Db 688 CTATGAGAGCTGCGGGGCAATTAAGAACAG-TACCGCTGAGAGGCTGCAAT-STCCG 631

Qy 181 CTGCTTCGCGCAGAGAGAAATCTCCCTGCTGCTGCTCAAGAGTGTGTTCTGAG 240

Db 630 CTGCTTCGCGCAGAGAGAAATCTCCCTGCTGCTGCTCAAGAGTGTGTTCTGAG 571

Qy 241 GGGGCTGTTGATGTGTGATGCTCTTCTGAGAGGCTTCATGCTTACTGATCCG 300

Db 570 GGGGCTGTTGATGTGTGATGCTCTTCTGAGAGGCTTCATGCTTACTGATCCG 511

Qy 301 GGTGCGACGAGAGAAACAGAGCGTGCCTGCGCACCGTCTGAGCTCCGAGATGACAA 360

Db 510 GGTGCGACGAGAGAAACAGAGCGTGCCTGCGCACCGTCTGAGCTCCGAGATGACAA 451

Qy 361 GAGCAGCTGTGAGAGAAACATATGCTCTGAGACCGCCCTGTGCTCAAGAGACT-GGG 419

Db 450 GAGCAGCTGTGAGAGAAACATATGCTCTGAGACCGCCCTGTGCTCAAGAGACTGGGG 391

Qy 420 AAGGAGGGGAGACTATGTGTGAGCTTTTAAATAGAGGATGACTGGATTTGAGT 479

Db 390 AAGGAGGGGAGACTATGTGTGAGCTTTTAAATAGAGGATGACTGGATTTGAGT 331

Qy 480 GATCATTAGAGCTGAGACTGTGTTCTCTGAGAGGTGAGACGCTGCTTCTGCTGAGCA 539

Db 330 GATCATTAGAGCTGAGACTGTGTTCTCTGAGAGGTGAGACGCTGCTTCTGCTGAGCA 271

Qy 540 GGGATGGGTTTCTTTGGAATACTCTTAGAGAGCTCTCTGCAATGGCTCGAGTCTGG 599

Db 270 GGGATGGGTTTCTTTGGAATACTCTTAGAGAGCTCTCTGCAATGGCTCGAGTCTGG 211

Qy 600 CAGAGCCCGAGTGTCTCTGCTGATGATTTCTTCCACAGGATGATTTCTTT 659

Db 210 CAGAGCCCGAGTGTCTCTGCTGATGATTTCTTCCACAGGATGATTTCTTT 151

Qy 660 GCTTAGTGAATTCATGCTCTTTTCTCATCAGAGAGTATGTGATGATGTTCTT 719

Db 150 GCTTAGTGAATTCATGCTCTTTTCTCATCAGAGAGTATGTGATGATGTTCTT 91

Qy 720 TTTGTTGTCGATTTATGATTTTATTAAGTATTAACAAAAGCTTTTATTAGATTCG 779

Db 90 TTTGTTGTCGATTTATGATTTTATTAAGTATTAACAAAAGCTTTTATTAGATTCG 31

Qy 780 AAAGAGAGAAAGTAAATGTAAGTTTA 808

Db 30 AAAGAGAGAAAGTAAATGTAAGTTTA 2

RESULT 11
AL575764/c 1062 bp mRNA linear EST 16-FEB-2001
LOCUS AL575764 LTI_NFL006_P12 Homo sapiens cDNA clone CS0D1070Y115 3
DEFINITION prime, mRNA sequence.
ACCESSION AL575764
VERSION AL575764.1 GI:12937249
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1062)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1..1062
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1070Y115"
/clone_lib="LTI_NFL006_P12"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
 Email : filang@lifetech.com URL :
 http://fulllength.invitrogen.com"

BASE COUNT 277 a 300 c 252 g 227 t 6 others
 ORIGIN
 Query Match 86.2%; Score 752.8; DB 9; Length 1062;
 Best Local Similarity 98.5%; Pred. No. 1.5e-127;
 Matches 786; Conservative 5; Mismatches 4; Indels 3; Gaps 3;

QY 1 CTCGAGGATATGTTCACTAATGAAAGAAATACCTGACCGCCAGCGAGTCACTGGGCGCTTG 60
 DB 804 CTCGAGGATATGTTCACTAATGAAAGAAATACCTGACCGCCAGCGAGTCACTGGGCGCTTG 745
 QY 61 CCGTCATCTCTCCACGCTGTAATGAGAGAGAACTCTGCAATTAATCTTAT 120
 DB 744 CCGTCATCTCTCCACGCTGTAATGAGAGAGAACTCTGCAATTAATCTTAT 685
 DB 684 CTATGAGAGCTGCCGGGCAATAGAAC-ATACCGCTCTGAGAGAGCTTGAT-CTCCG 627
 QY 181 CTGCTCCGCGACAGAGAAATCTCCCTGCTGCTCAAGAGTGTGTCTGCG 240
 DB 628 CTGCTCCGCGACAGAGAAATCTCCCTGCTGCTCAAGAGTGTGTCTGCG 567
 QY 241 GGGGCTGTCTGATGATGATCTCTCTGAGAGCTCCAGATGATCTGATCG 300
 DB 566 GGGGCTGTCTGATGATGATCTCTCTGAGAGCTCCAGATGATCTGATCG 507
 QY 301 GGTGACAGAGAAACAGAGAGCTGCGCTGAGAGCTGAGAGCTCCGAGATGACAA 360
 DB 506 GGTGACAGAGAAACAGAGAGCTGCGCTGAGAGCTGAGAGCTCCGAGATGACAA 447
 QY 361 GAGAGAGTGTGAAACACATATGCTGTGACCGCTGTCGCAAGAGACT-GGG 419
 DB 446 GAGAGAGTGTGAAACACATATGCTGTGACCGCTGTCGCAAGAGACTGAGG 387
 QY 420 AAGGAGGAGAGAACTATGATGAGAGCTTTTAAATAGAGAGATTAACCTGATTTAGT 479
 DB 386 AAGGAGGAGAGAACTATGATGAGAGCTTTTAAATAGAGAGATTAACCTGATTTAGT 327
 QY 480 GATCATTAAGGCTGAGGCTGTTCTCTGAGAGTAGAGAGCTGTTCTGATGAGCA 539
 DB 326 GATCATTAAGGCTGAGGCTGTTCTCTGAGAGTAGAGAGCTGTTCTGATGAGCA 267
 QY 540 GGGATGGTTGCTTTGGAATCTCTAGAGAGCTCTCTGCAATGAGCTGAGCTGG 599
 DB 266 GGGATGGTTGCTTTGGAATCTCTAGAGAGCTCTCTGCAATGAGCTGAGCTGG 207
 QY 600 CAGCAGCCCGGAGTGTCTCTGCGTGAATTTCTTCTCTGAGTAGAGTTTCTT 659
 DB 206 CAGCAGCCCGGAGTGTCTCTGCGTGAATTTCTTCTCTGAGTAGAGTTTCTT 147
 QY 660 GCTTATGTAATTCATGCTCTTCTCATCAGAGAGTAGATGTAATGTTTCT 719
 DB 146 GCTTATGTAATTCATGCTCTTCTCATCAGAGAGTAGATGTAATGTTTCT 87
 QY 720 TTTGTTGTCTGATTAATGTTTAAATAGTAATTAAGTAATTAATGATTTCTG 779
 DB 86 TTTGTTGTCTGATTAATGTTTAAATAGTAATTAAGTAATTAATGATTTCTG 27
 QY 780 AAAAGAGAGAAATTAAT 797
 DB 26 AAAAGAGAGAAATTAAT 9

RESULT 12
 AL560664/c
 LOCUS AL560664 964 bp mRNA linear EST 16-FEB-2001
 DEFINITION AL560664 LTI NFL010, BC2 Homo sapiens cDNA clone CS0D.U003YCI1 3
 pt1me, mRNA sequence.
 ACCESSION AL560664

VERSION AL560664.1 GI:12907346
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 964)
 AUTHORS Li, W.-B., Gruber, C., Jessee, J., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 Location/Qualifiers

FEATURES
 source
 1. 964
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="OSD.L003YCI17"
 /clone_1fb="LTI_NFL010_BC2"
 /sex="male"
 /tissue_type="B cells from Burkitt lymphoma"
 /note="Vector: pCMVSPORT 6; Site 1: NotI, 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : filang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 255 a 274 c 233 g 199 t 3 others
 ORIGIN

Query Match 85.7%; Score 748; DB 9; Length 964;
 Best Local Similarity 99.1%; Pred. No. 1.2e-126;
 Matches 770; Conservative 2; Mismatches 3; Indels 2; Gaps 2;

QY 1 CTCGAGGATATGTTCACTAATGAAAGAAATACCTGACCGCCAGCGAGTCACTGGGCGCTTG 60
 DB 787 CTCGAGGATATGTTCACTAATGAAAGAAATACCTGACCGCCAGCGAGTCACTGGGCGCTTG 728
 QY 61 CCGTCATCTCTCCACGCTGTAATGAGAGAGAACTCTGCAATTAATCTTAT 120
 DB 727 CCGTCATCTCTCCACGCTGTAATGAGAGAGAACTCTGCAATTAATCTTAT 668
 QY 121 CTATGAGAGCTGCCGGGCAATAG-AAACAGTACCGCTCTGAGAGAGCTGATGCTCC 179
 DB 667 CTATGAGAGCTGCCGGGCAATAG-AAACAGTACCGCTCTGAGAGAGCTGATGCTCC 608
 QY 180 GCTGCTTCGCGACAGAGAAATCTCCCTGCGCTTGGCTCAAAAGTGTGTCTGG 239
 DB 607 GCTGCTTCGCGACAGAGAAATCTCCCTGCGCTTGGCTCAAAAGTGTGTCTGG 548
 QY 240 CCGGCTCTTCGAGAGTGTGATCTCTCTGAGAGAGCTCCAGAGTGTGATCTGATGCC 299
 DB 547 CCGGCTCTTCGAGAGTGTGATCTCTCTGAGAGAGCTCCAGAGTGTGATCTGATGCC 488
 QY 300 GGGTGGCAGGAGAGAACAGAGAGCTGCGCAGCGCTGAGAGCTCCGAGATGACA 359
 DB 487 GGGTGGCAGGAGAGAACAGAGAGCTGCGCAGCGCTGAGAGCTCCGAGATGACA 428
 QY 427 AAGAGCGCTGTGTAAGAACACATATGCTGTGACCGCTCTGCAAGAGAGCTGG 368
 DB 419 GAAAGGAGGAGAGCAATATGCTGTGAGCTTTTAAATAGAGGATGATCGGATTTAG 478
 QY 367 GAAAGGAGGAGAGCAATATGCTGTGAGCTTTTAAATAGAGGATGATCGGATTTAG 308
 QY 479 TGATCATTAAGGCTGAGAGCTGTTCTCTGAGAGTGAAGAGCTGCTCTGCTGCTGGC 538

Db 307 TGACATTAAGGCGTAGAGTCTGTATTCTCTGAGGAGGTAGAGCGGCTCTTCCGAGCTGAGC 248

QY 539 AGGGATGGGTTTGCTTTGGAAATCCTCTAGAGAGCTCTCTCTGCAATGGCTGCAGACTG 598

Db 247 AGGGAAGGGTTTGTTTGGAAATCCCTTAGAGAGGCTCCYCTCGCATGGCTCGCAGACTG 188

QY 599 GCGAGAGCGCCGAGTGTTCCTCGGTATGCAATTTCTTTCCTCGAGGAGAGTTTCTT 658

Db 187 GCAGAGAGCCCGAGTGTTCCTCGGTATGCAATTTCTTTCCTCGAGGAGAGTTTCTT 128

QY 659 TGGTATGTGAATTCATTGCTCTCTTTCCTATCAGAGAGTATGTGAATCGTTTC 718

Db 127 TGGTATGTGAATTCATTGCTCTCTTTCCTATCAGAGAGTATGTGAATCGTTTC 68

QY 719 TTTTGTGTCTGATTTATAGTTTTTTTAAAGTAAACAAAGTTTTTATATAGAT 775

Db 67 TTTTGTGTCTGATTTATAGTTTTTTTAAAGTAAACAAAGTTTTTATATAGAT 11

RESULT 13	
AL582333/c	
LOCUS	859 bp mRNA linear EST 16-FEB-2001
DEFINITION	AL582333 LIT_NFL010.BC2 Homo sapiens cDNA clone CS0DL006Y120 3 prime, RNA sequence.
ACCESSION	AL582333
VERSION	AL582333.1 GI:12950213
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE	1 (bases 1 to 859)
AUTHORS	Li, W.-B., Gruber, C., Jesse, J. and Polyes, D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	Contact: Genoscope

Genoscope - Centre National de Séquençage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr.

```
FEATURES
source      Location/Qualifiers
1. .859
```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/cdone="CS0DL006Y20"
/cdome="LTT_NFL010_BC2"
/sex="male"

```

BASE COUNT 230 a 240 c 204 g 178 t 7 others

Query Match	85.7%;	Score 747.8;	DB 9;	Length 859;
Best Local Similarity	98.6%;	Pred. No. 1.4e-126;		
Matches 790; Conservative	5;	Mismatches	2;	Incls 4;
				Gaps 4

QY 1 CTCGAGCATATGTTCACTATGAGATACTGCACCGCCCAACGACAGTTCATCTGGCCCTTG 60

Db 800 CTCACGCAATATGTTCAACTATGAGAACTACTGACACCGCCACAGCAGTACT -GGCCTTG 742

61 CCGTGCATCCTTCCACGCGTGTACTTTTGAAGTGAAGGAATCCTGCAATAAATTGAT 120

741 CCGTGCATCCTTCCACGCTGGTACTTTGACGTGGAGAGGAACTCTGCATTAATCTCAT 682

121 CTATGAGGCTGCGCGGGCAATAA-GAA CAGCTACCGCTCTGA-GGAGGCTGCATGCTC 178

Db 681 CTTTGAAGGCTGCCGGGCGAATTAAGAAACAGTACCGCTCTTAGGAGAGCCTGCATTGCT 622

QY 179 CGCTCTTCGCGCAGCAGAGAAATCCTCCCTGCCCCCTTGGCTCAAGAGTGGTGTCTG 238

Db 621 CGGTCTTCGCGCAGCAGAGAAATCCTCCCTGCCCCCTTGGCTCAAGAGTGGTGTCTG 562

QY 239 GCGGGGCTGTTCGTGATGTTGATCTCTTCTTCTGGAGACCTCCATGTGTTCACCTGATC 298

Db 561 GCGGGGCTGTTCGTGATGTTGATCTCTTCTTCTGGAGACCTCCATGTGTTCACCTGATC 502

QY 299 CGGCTGACACCGAGAAACAGAGAGCGTCCCTGCGCACCGTCTGACAGCTCCGAGATGAC 358

Db 501 CGGCTGACACCGAGAAACAGAGAGCGTCCCTGCGCACCGTCTGACAGCTCCGAGATGAC 442

QY 359 AAGGAGCAGCTGTGAAAGAACATATGTCTCTGACCGCCCTGTGCGCAAGAGACT-G 417

Db 441 AAGGAGCAGCTGTGAAAGAACATATGTCTCTGACCGCCCTGTGCGCAAGAGAGTGG 382

418 GGAAGGGAGGGGACATATGTGTGAGCTTTTAAATAGAGGGATTGACTCGGATTTGA 477

Db 381 GGAAGGAGGGGAGACTAAGTGTGAGCTTTTAAAGAGGGATTGACTGGGATTTCG 322

24 478 GTGATCATTAGGCTGAGCTCTGTTTCTCTGGAGGTAGGACGGCTCTTCTGTGG 537

db 321 GTGATCATTAGGCTGAGCTCTGTTCTCTGAGAGTACGACCGCTTCTCTGGTCTGG 262

538 CAGGATGGGTTTGCCTTTGGAAATCCTCTAGGAGGCTCCTCCTCGCATGGCCTGCAGTCT 597

Db 261 CAGGATGGGTTTCTTTGGAAATCTCTAGGAGGCTCCCCCGCATGGCCTGCAGTT 202

598 GGCAGAGCCCCGAGTGTTCCTCGCTGATCGATTCTTCTCCACGAGGTAGATTTCCT 657

Db 201 GGCAGCAGCCCGAGTGTTCCTCGGTACGATTTCTTCCCTCAGGTAGATTTCT 142

658 TTGCTATGTTGAATTCATGCGCTTTTCTCATCAGAGAGTATGTTGCAATCGTTT 717

Db 141 TTGCTATGTTGAATTCATTGCCCTTTTTCATCAGAGAAGTATGKTGGAATCGT 82

718 CTTTGCTTGCATTAATGCTTTTAAAGTATAACAAAGCTTTTATATAGCATTG 777

Db

81 CTTTGTGTCGATTTATGTTTTTTAAGATAAACAAMTTTTTATTAGATTC 22

0v 778 TGAAGTAGGAAAGTAAATG 798

21 TCAAGCAAGGAAGTAAATG 1

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

RESULT 14

AL571794	990 bp	mRNA	linear	EST 16-FEB-2001
----------	--------	------	--------	-----------------

DEFINITION
AL511194 L11 NFUB6_P12 homo sapiens cDNA clone CSOD10310C1/3
prime, mRNA sequence.

ACCESSION	AL571794
VERSION	AL571794.1
	GI:12929445

KEYWORDS	EST.
SOURCE	human.

ORGANISM Homo sapiens
 Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa;

REFERENCE 1 (bases 1 to 990)
Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.

AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France

FRATRES
 Email : secref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 Location/Qualifiers

```

source
1. .990
/organism="Homo sapiens"

```

```

/organisms=Juncus sagittalis
/db_xref="taxon:9606"
/olname="JCSN07030V017"

```

/CI0101E CS001030VCL17

```

/clone_lib="LTI_NF1006_P12"
/issue_type="placenta"
/notes="Vector: PCWSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the PCWSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT      258 a      282 c      244 g      203 t      3 others
ORIGIN

```

```

Query Match      85.0%; Score 742.2; DB 9; Length 990;
Best Local Similarity 99.6%; Pred. No. 1,4e-125;
Matches 764; Conservative 1; Mismatches 0; Indels 2; Gaps 2;

```

```

QY 1 CTCGAGGATGATGTTCACTAATGAAGATCTGACCGCCCAAGCAGTCACTGGGCTTTG 60
DB 767 CTCGAGGATGATGTTCACTAATGAAGATCTGACCGCCCAAGCAGTCACTGGGCTTTG 708
QY 61 CCGTGCACTCTTCCACGCTGGTACTTGAAGTGAAGAGAACTCTGCAATTAATTCTTCA 120
DB 707 CCGTGCACTCTTCCACGCTGGTACTTGAAGTGAAGAGAACTCTGCAATTAATTCTTCA 648
QY 121 CTATGAGAGGCTGCGGGGCAATTAAGAA-CAGCTACCGCTCTGAGAGAGGCTTGATGCTCC 179
DB 647 CTATGAGAGGCTGCGGGGCAATTAAGAA-CAGCTACCGCTCTGAGAGAGGCTTGATGCTCC 588
QY 180 GCTGCTTCCGACGAGAGAGATCTCCCTGCTGGCTCAAGAGTGGTGGTCTG 239
DB 587 GCTGCTTCCGACGAGAGAGATCTCCCTGCTGGCTCAAGAGTGGTGGTCTG 528
QY 240 CGGAGCTGTTTCTGATGATGATGATCTCTCTCTGAGACCTCCATGCTTCACTGATCC 299
DB 527 CGGAGCTGTTTCTGATGATGATGATCTCTCTCTGAGACCTCCATGCTTCACTGATCC 468
QY 300 GGGTGGCAGGAGAAACCAGAGAGCTGCTGGCAACGCTCTGAGCTCCGAGATGACA 359
DB 467 GGGTGGCAGGAGAAACCAGAGAGCTGCTGGCAACGCTCTGAGCTCCGAGATGACA 408
QY 360 AGGAGCAGCTGGTGAAGAAACATATGCTGTGACCGCCCTGCGCAAGAGACT-GG 418
DB 407 AGGAGCAGCTGGTGAAGAAACATATGCTGTGACCGCCCTGCGCAAGAGACTGGG 348
QY 419 GAAAGGAGGAGAGACTATGTGTGAGCTTTTAAATAGAGGATTAAGCTCGATTGAG 478
DB 347 GAAAGGAGGAGAGACTATGTGTGAGCTTTTAAATAGAGGATTAAGCTCGATTGAG 288
QY 479 TGATCATTAGGGCTAGAGTCTGTTCTCTGAGAGGTAGGACGGCTGCTTCTGCTGGC 538
DB 287 TGATCATTAGGGCTAGAGTCTGTTCTCTGAGAGGTAGGACGGCTGCTTCTGCTGGC 228
QY 539 AGGAGATGGGTTGCTTGGAAATCCTTAGAGGCTCCCTCCGAGAGCCTGCAATTCG 598
DB 227 AGGAGATGGGTTGCTTGGAAATCCTTAGAGGCTCCCTCCGAGAGCCTGCAATTCG 168
QY 599 GCAAGACCCCGAGATGTTTCTCTGATGATGATTTCTTCTCCAGGAGAGTTTCTT 658
DB 167 GCAAGACCCCGAGATGTTTCTCTGATGATGATTTCTTCTCCAGGAGAGTTTCTT 108
QY 659 TGCCTATGTTGAATCCATTCCTCTTTTTCATCAGAGATGATGTTGATTC 718
DB 107 TGCCTATGTTGAATCCATTCCTCTTTTCATCAGAGATGATGTTGATTC 48
QY 719 TTTTGTGCTGATTAATGTTTTTAAAGTAAAGAAAGTTT 765
DB 47 TTTTGTGCTGATTAATGTTTTTAAAGTAAAGAAAGTTT 1

```

RESULT 15

```

BM977432/c
LOCUS      BM977432      750 bp      mRNA      linear      EST 21-MAR-2002
DEFINITION      UI-CF-EN1-see-h-15-0-UI-s1 UI-CF-EN1 Homo sapiens cDNA clone
ACCESSION      UI-CF-EN1-see-h-15-0-UI 3', mRNA sequence.
VERSION      BM977432.1 GI:19595843
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 750)
AUTHORS      Bonaldo,M.F., Lennon,G., and Soares,M.B.
TITLE      Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL      Genome Res. 6 (9), 791-806 (1996)
MEDLINE      97044477
COMMENT      Contact: McCray, PB
              McCray Lab
              University of Iowa
              2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
              Tel: 319 356 4866
              Fax: 319 356 7171
              Email: paul-mccray@uiowa.edu
              Tissue Procurement: Dr. M. J. Welsh, University of Iowa
              CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
              CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
              DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
              Clone Distribution: Researchers may obtain clones from Research
              Genetics (www.resgen.com).
              Seq primer: M13 FORWARD
              POLYA=yes.
FEATURES
Source
Location/Qualifiers
1..750
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-CF-EN1-see-h-15-0-UI"
/clone_lib="UI-CF-EN1"
/tissue_type="Primary Lung Cystic Fibrosis Epithelial
Cells"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/notes="Organ: Lung; Vector: pRTT3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-CF-EN1 is a normalized cDNA library containing the
following tissue(s): Primary Lung Cystic Fibrosis
Epithelial Cells. The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was ligated to an EcoR I adaptor, digested with Not
I, and cloned directionally into pRTT3-Pac vector. The
oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is CTGCTCAGCT.
TAG_Lib=UI-CF-EN1
TAG_Tissue=Human Lung Epithelial Cell lines untreated LPS
6hr to LPS 24h
TAG_SBO=CTGCTCAGCT"

```

```

BASE COUNT      213 a      209 c      160 g      168 t
ORIGIN
Query Match      84.5%; Score 738; DB 14; Length 750;
Best Local Similarity 99.9%; Pred. No. 9.2e-125;
Matches 749; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```

```

QY 125 GAGAGCTGCCGGGCAATTAAGACAGTACCGCTCTGAGAGAGCCTGATGCTCCGCTGC 184
DB 750 GAGAGCTGCCGGGCAATTAAGACAGTACCGCTCTGAGAGAGCCTGATGCTCCGCTGC 691
QY 185 TTCCGACGAGAGAGATCTCCCTGCTGCTCAAGAGTGGTGGTCTTCTGAGGCGG 244

```

Db 690 TTCCGCCACAGAGAACTCCCTGCCCCCTGGCTGAGCTCAAGGTGTGTTCTGGCGGG 631
QY 245 CTGTTGCTGATGCTGTGATCTCTCTCTGGAGCCTCAATGCTACTGATCCGGGTG 304
Db 630 CTGTGTGATGTGTGTGATCTCTCTCTGGAGCCTCAATGCTACTGATCCGGGTG 571
QY 305 GCACGGAGGAAACAGAGACCTGCTCCGACCGTCTGAGCTCCGAGATGACAAAGAG 364
Db 570 GCACGGAGGAAACAGAGACCTGCTCCGACCGTCTGAGCTCCGAGATGACAAAGAG 511
QY 365 CAGCTGTGAGAACACATATGCTGTGACCGCCCTGTGCGCAAGAGACT-GGGAGG 423
Db 510 CAGCTGTGAGAACACATATGCTGTGACCGCCCTGTGCGCAAGAGACTGGGAGG 451
QY 424 GAGGGAGACTGTGTGAGCTTTTAAATAGAGGATTGACTCGGATTTGAGTGATC 483
Db 450 GAGGGAGACTGTGTGAGCTTTTAAATAGAGGATTGACTCGGATTTGAGTGATC 391
QY 484 ATTAGGGCTGAGTCTGTTTCTGGAGGAGTAGACGGCTGCTTCTGTGCTGGCAGGGA 543
Db 390 ATTAGGGCTGAGTCTGTTTCTGGAGGAGTAGACGGCTGCTTCTGTGCTGGCAGGGA 331
QY 544 TGGGTTTGGCTTGGAAATCCTTAGAGGCTCCTCTCGATGGCTGCAGTGCGCAGC 603
Db 330 TGGGTTTGGCTTGGAAATCCTTAGAGGCTCCTCTCGATGGCTGCAGTGCGCAGC 271
QY 604 AGCCCCGAGTGTTCCTCGCTGATCGATTCTTTCCTCCAGGTAGAGTTTCTTGGCTT 663
Db 270 AGCCCCGAGTGTTCCTCGCTGATCGATTCTTTCCTCCAGGTAGAGTTTCTTGGCTT 211
QY 664 ATGTGAATTCATGCTCCTTTCTCATGACAGAAAGTAGATGTAATGTTCTTTTG 723
Db 210 ATGTGAATTCATGCTCCTTTCTCATGACAGAAAGTAGATGTAATGTTCTTTTG 151
QY 724 TTGTCTGATTTATGTTTTTAAAGTAAACAAAGTTTTTATTAGCATCTGAAAG 783
Db 150 TTGTCTGATTTATGTTTTTAAAGTAAACAAAGTTTTTATTAGCATCTGAAAG 91
QY 784 AAGGAAAGTAAATGTACAGTTTAATAAAAAGGGCCTTCCCTTTAGAAATAATTCA 843
Db 90 AAGGAAAGTAAATGTACAGTTTAATAAAAAGGGCCTTCCCTTTAGAAATAATTCA 31
QY 844 GCATGTCTTCAAAAAAAAAAAAAAAAAAAAA 873
Db 30 GCATGTCTTCAAAAAAAAAAAAAAAAAAAAA 1

Search completed: January 10, 2003, 03:20:49
Job time : 1883.25 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 10, 2003, 01:44:31 ; Search time 714.023 Seconds

(without alignments)
13246.642 Million cell updates/sec

Title: US-09-825-682A-57
Perfect score: 325
Sequence: 1 aaagggcgccgagcgccct.....gtacaaaaaaaaaaaaa 325

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenBml:*

1: gb ba:*

2: gb hrg:*

3: gb in:*

4: gb om:*

5: gb ov:*

6: gb pat:*

7: gb ph:*

8: gb pl:*

9: gb pr:*

10: gb ro:*

11: gb sts:*

12: gb sy:*

13: gb un:*

14: gb vi:*

15: em ba:*

16: em fun:*

17: em hum:*

18: em in:*

19: em mu:*

20: em on:*

21: em or:*

22: em ov:*

23: em pat:*

24: em ph:*

25: em pl:*

26: em ro:*

27: em sts:*

28: em un:*

29: em vi:*

30: em hrg hum:*

31: em hrg inv:*

32: em hrg other:*

33: em hrg mus:*

34: em hrg pin:*

35: em hrg rod:*

36: em hrg mam:*

37: em hrg vrt:*

38: em sy:*

39: em hrgo hum:*

40: em hrgo mus:*

41: em hrgo other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	320.4	98.6	2484	6	AX454042
2	320.4	98.6	2484	9	BC008765
3	307.4	94.6	2430	6	AR068043
4	303.4	93.4	1610	6	AX017423
5	303.4	93.4	53135	9	AC104792
6	303.4	93.4	148773	2	AC041029
7	289.4	89.0	2402	9	HUMSYN
8	285.8	87.9	4797	6	AX36413
9	285.8	87.9	4797	6	AX365741
10	285.8	87.9	4797	6	AX411295
11	285.8	87.9	4797	9	HSYNDIGN
12	181.4	55.8	425	6	AX454046
13	144.4	44.4	492	6	AX368331
14	130	40.0	200548	2	AL672026
15	129.8	39.9	194985	10	AC002406
16	129.8	39.9	200548	2	AL672026
17	126.2	38.8	2410	10	S61865
18	120.4	37.0	2396	10	RATSYNDECA
19	120.4	37.0	176793	2	AC123473
20	113	34.8	26700	6	AR068044
21	113	34.8	26700	6	191962
22	112.6	34.6	33934	10	MMSYNDEIA
23	112.6	34.6	2432	6	117250
24	112.6	34.6	2432	10	MMSYNDE
25	111.2	34.5	2985	10	HAMPSR
26	111.4	34.3	2513	10	BC010560
27	105.8	32.6	133499	2	AC094856
28	49	15.1	156533	2	AC117070
29	47.4	14.6	152409	2	PFMAL1P1
30	44.8	13.8	123280	2	AC117076
31	44.6	13.7	3565	3	AC104041
32	43.4	13.4	14433	3	AB001369
33	43.4	13.4	121181	9	AC092821
34	43.4	13.4	180685	9	AC006432
35	43.4	13.4	240342	2	AC092862
36	43.2	13.3	180231	9	AC009161
37	43.2	13.3	318221	2	PFMAL1P3
38	42.8	13.2	87180	8	AB023044
39	42.8	13.2	162200	9	AC012156
40	42.8	13.2	205952	9	AC024940
41	42.6	13.1	868	9	AK024813
42	42.6	13.1	2476	9	BC027954
43	42.6	13.1	2518	3	AF309947
44	42	12.9	1853	9	BC028120
45	41.8	12.9	3787	10	BC006627

ALIGNMENTS

RESULT 1
LOCUS AX454042
DEFINITION Sequence 18 from Patent WO0198539.
ACCESSION AX454042
VERSION AX454042.1 GI:21713662
KEYWORDS
SOURCE
ORGANISM human.
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Mitsuhashi,M., Kambara,H., Matsumaga,H. and Kawamura,M.
TITLE Gene markers for lung cancer
JOURNAL Patent: WO 0198539-A 18 27-DEC-2001;

TITLE Suppression of tumor cell growth by syndecan-1 ectodomain
JOURNAL Patent: US 5851993-A 3 22-DEC-1998;
FEATURES Location/Qualifiers
SOURCE 1. 2430
/organism="unknown"
BASE COUNT 490 a 735 c 712 g 492 t 1 others
ORIGIN
Query Match 94.6%; Score 307.4; DB 6; Length 2430;
Best Local Similarity 99.1%; Pred. No. 2,4e-66;
Matches 319; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 4 GAGGGCGAGGGGCTGAGATCCTCTGACAGCAGCCGCTCTGAGCGCG 63
DB 2100 GAGGGCGAGGGGCTGAGATCCTCTGACAGCAGCCGCTCTGAGCGCG 2158
QY 64 TCTCCAGGGGCTGCTTCTCTGGAATGACAGAGGGGTGTCTTGGGCAAGCTGGCTCT 123
DB 2159 TCTCCAGGGGCTGCTTCTCTGGAATGACAGAGGGGTGTCTTGGGCAAGCTGGCTCT 2218
QY 124 GAGGGCTCCATCCAGGCGAGGTTCTCCGTTAGCTCTGTGCCCCACCTGGGCGCTTG 183
DB 2219 GAGGGCTCCATCCAGGCGAGGTTCTCCGTTAGCTCTGTGCCCCACCTGGGCGCTTG 2278
QY 184 GGCTGGATCAGGAATATTTTCCAAAGATGATAGCTTTTGGCTTTGGCAAACTCTAC 243
DB 2279 GGCTGGATCAGGAATATTTTCCAAAGATGATAGCTTTTGGCTTTGGCAAACTCTAC 2338
QY 244 TTAATCCATGGGTTTCTCTGTACAGTAGATTTTCCAAATGTAATTAATTAATA 303
DB 2339 TTAATCCATGGGTTTCTCTGTACAGTAGATTTTCCAAATGTAATTAATTAATA 2398
QY 304 AAGTAAAAAAGGTTTCTCTGTACAGTAGATTTTCCAAATGTAATTAATTAATA 325
DB 2399 AAGTAAAAAAGGTTTCTCTGTACAGTAGATTTTCCAAATGTAATTAATTAATA 2420
RESULT 4
LOCUS AX017423 1610 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 202 from Patent WO9947669.
ACCESSION AX017423
VERSION AX017423.1 GI:10042244
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1610)
AUTHORS Schmitt,A., Specht,T., Dahl,E., Hinemann,B., Rosenthal,A. and
Pillarsky,C.
TITLE Human nucleic acid sequences from tissue of breast tumors
JOURNAL Patent: WO 9947669-A 202 23-SEP-1999;
SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
BERND (DE); ROSENTHAL ANDRE (DE); METZGEN GRS FUER GENOMFORSCHUN
(DE); PILLARSKY CHRISTIAN (DE)
FEATURES Location/Qualifiers
SOURCE 1. 1610
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 312 a 409 c 456 g 433 t
ORIGIN
Query Match 93.4%; Score 303.4; DB 6; Length 1610;
Best Local Similarity 99.7%; Pred. No. 2,4e-65;
Matches 304; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 GAGGGCGAGGGGCTGAGATCCTCTGACAGCAGCCGCTCTGAGCGCG 63
DB 601 GAGGGCGAGGGGCTGAGATCCTCTGACAGCAGCCGCTCTGAGCGCG 660
QY 64 TCTCCAGGGGCTGCTTCTCTGGAATGACAGAGGGGTGTCTTGGGCAAGCTGGCTCT 123

DB 661 TCTCCAGGGGCTGCTTCTCTGGAATGACAGAGGGGTGTCTTGGGCAAGCTGGCTCT 720
QY 124 GAGGGCTCCATCCAGGCGAGGTTCTCCGTTAGCTCTGTGCCCCACCTGGGCGCTTG 183
DB 721 GAGGGCTCCATCCAGGCGAGGTTCTCCGTTAGCTCTGTGCCCCACCTGGGCGCTTG 780
QY 184 GGCTGGATCAGGAATATTTTCCAAAGATGATAGCTTTTGGCTTTGGCAAACTCTAC 243
DB 781 GGCTGGATCAGGAATATTTTCCAAAGATGATAGCTTTTGGCTTTGGCAAACTCTAC 840
QY 244 TTAATCCATGGGTTTCTCTGTACAGTAGATTTTCCAAATGTAATTAATTAATA 303
DB 841 TTAATCCATGGGTTTCTCTGTACAGTAGATTTTCCAAATGTAATTAATTAATA 900
QY 304 AAGTA 308
DB 901 AAGTA 905

RESULT 5

LOCUS AC104792/c 5313 bp DNA linear PRI 01-JUN-2002
DEFINITION Homo sapiens BAC clone RP11-202B22 from 2, complete sequence.
ACCESSION AC104792
VERSION AC104792.5 GI:20800378
KEYWORDS
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 5313)
AUTHORS Sulston,J.B. and Waterston,R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PubMed 9847074

REFERENCE

2 (bases 1 to 5313)
AUTHORS Abbott,S., Kozlowski,A. and Dignan,G.
TITLE The sequence of Homo sapiens BAC clone RP11-202B22
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 5313)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (21-DEC-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

REFERENCE

4 (bases 1 to 5313)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (11-APR-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

REFERENCE

5 (bases 1 to 5313)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (15-MAY-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

REFERENCE

6 (bases 1 to 5313)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (01-JUN-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On May 15, 2002 this sequence version replaced gi:20136955.

COMMENT

----- Genome Center
Center: Washington University Genome Sequencing Center
Genetic code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sepiens@wustl.edu
----- Summary Statistics
Center project name: H_NH0202B22

NOTE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCT-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, E., Tateo, M., Caranese, J.J., and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP13-496P22; the clone sequenced to the right is RP11-327N17, 2000 bp overlap. Actual start of this clone is at base position 1 of RP11-202B22; actual end is at base position 98942 of RP11-327N17.

Polymorphisms exist between AC104792 and AC116171. Data from AC116171 was used to finish AC104792.

FEATURES

source

1..53135
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"

/map="2"

/clone="RP11-202B22"

/clone_id="RPCT-11"

184..251

/rpt_family="L2"

1876..1926

/rpt_family="MIR"

2632..2785

/rpt_family="MIR"

3411..3447

/rpt_family="MIR"

4149..4284

/rpt_family="CA)n"

4410..4728

/rpt_family="MIR"

5197..5405

/rpt_family="L2"

5368..5573

/rpt_family="MIR"

6069..6171

/rpt_family="MIR"

7102..7181

/rpt_family="MIR"

7381..7519

/rpt_family="MIR"

8154..8517

/rpt_family="Alu"

8628..8947

/rpt_family="L1"

9450..9662

/rpt_family="Alu"

/rpt_family="L2"

repeat_region

9666..9795

repeat_region

9796..10090

repeat_region

10091..10207

repeat_region

10383..10521

repeat_region

10624..10907

repeat_region

10970..11376

repeat_region

12019..12609

repeat_region

13239..13287

repeat_region

19556..19649

repeat_region

19736..19816

repeat_region

20416..20502

repeat_region

21437..21656

repeat_region

22223..22512

repeat_region

23163..23280

repeat_region

23985..24117

repeat_region

24848..25084

repeat_region

29614..29827

repeat_region

30410..30602

repeat_region

32299..32459

repeat_region

32945..33052

repeat_region

33301..33478

repeat_region

33506..33652

repeat_region

34099..34137

repeat_region

35124..35177

repeat_region

35207..35352

repeat_region

35451..35592

repeat_region

35966..36071

repeat_region

36872..36894

repeat_region

37331..37432

repeat_region

38111..40789

repeat_region

39015..39091

repeat_region

39117..39152

repeat_region

39233..39263

repeat_region

41509..41715

repeat_region

41716..41945

repeat_region

41946..42234

repeat_region

42271..42758

repeat_region

42271..42758

repeat_region

42271..42758

repeat_region

42271..42758

repeat_region

42271..42758

repeat_region

42271..42758

repeat_region

42271..42758

repeat_region

42271..42758

repeat_region

42271..42758

repeat_region

42271..42758

repeat_region

42271..42758

repeat_region

42271..42758

repeat_region

42271..42758

repeat_region

42271..42758

repeat_region

42271..42758

repeat_region

42271..42758

repeat_region

42271..42758

repeat_region

42271..42758

repeat_region

42271..42758

repeat_region

42271..42758

repeat_region

42271..42758

repeat_region

42271..42758

repeat_region

42271..42758

repeat_region

42271..42758

repeat_region

42271..42758

repeat_region

42271..42758

repeat_region

42271..42758

repeat_region

42271..42758

repeat_region

42271..42758

repeat_region

42271..42758


```

misc_feature      54026..61888
                  /note="assembly_fragment"
misc_feature      61989..69410
                  /note="assembly_fragment"
misc_feature      69511..76251
                  /note="assembly_fragment"
misc_feature      76352..84563
                  /note="assembly_fragment"
misc_feature      84664..102440
                  /note="assembly_fragment"
misc_feature      102541..123487
                  /note="assembly_fragment"
misc_feature      123588..146373
                  /note="assembly_fragment"
misc_feature      146474..149773
                  /note="assembly_fragment"
                  clone_end:77
                  vector_side:right"
BASE COUNT      40370 a 32119 c 32522 g 43660 t 1102 others
ORIGIN

```

```

Query Match      93.4%; Score 303.4; DB 2; Length 149773;
Best Local Similarity 99.7%; Pred. No. 2.7e-65;
Matches 304; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      4  GAGGGCGGAGGGGCTGAGATCCTCTGACAGACACGCCCTCTGCTGAGCGCG 63
Db      33395 GAGGGCGGAGGGGCTGAGATCCTCTGACAGACAGCCCGTCTGAGCGCG 33354

QY      64  TCTCCAGGGGCTGCTTCTCTGAAATTGACAGGGGTGCTTGGGAGAGTGGCTCT 123
Db      33355 TCTCCAGGGGCTGCTTCTCTGAAATTGACAGGGGTGCTTGGGAGAGTGGCTCT 33414

QY      124 GAGCGCTCCATCCAGGCGAGGTCTCCGTTAGCTCTGAGGCCACCTGGGCGCTG 183
Db      33415 GAGCGCTCCATCCAGGCGAGGTCTCCGTTAGCTCTGAGGCCACCTGGGCGCTG 33474

QY      184 GGTGGATCAGGAATATTTTCCAAAGATGATAGTCTTTTGGCAAACTCTAC 243
Db      33475 GGTGGATCAGGAATATTTTCCAAAGATGATAGTCTTTTGGCAAACTCTAC 33534

QY      244 TTAATCAATGGGTTTTCTCTGACAGTAGATTTTCCAAATGTAAATACTTAATATA 303
Db      33535 TTAATCAATGGGTTTTCTCTGACAGTAGATTTTCCAAATGTAAATACTTAATATA 33594

QY      304 AAGTA 308
Db      33595 AAGTA 33599

```

```

RESULT 7
HUMSYN      HUMSYN      2402 bp      mRNA      linear      PRI 13-JAN-1995
LOCUS      Human syndecan mRNA, complete cds.
DEFINITION      J05392
ACCESSION      J05392.1 GI:338633
VERSION      Integral membrane protein; syndecan.
KEYWORDS      Human breast cell line HBL-100, cDNA to mRNA, clones hsyn(4,pr7).
SOURCE      Homo sapiens
ORGANISM      Homo sapiens
REFERENCE      1 (bases 1 to 2402)
AUTHORS      Mamiya, T.; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE      Mamiya, T., Jaakkola, P., Arvilommi, A.M. and Jalkanen, M.
JOURNAL      Integral membrane proteoglycans
MEDLINE      J. Biol. Chem. 265 (12), 6884-6889 (1990)
PUBMED      90216719
COMMENT      Sequence of human syndecan indicates a novel gene family of
FEATURES      Draft entry and printed sequence for [1] kindly submitted by
SOURCE      M.Mali, 13-FEB-1990, for release after publication.
              Location/Qualifiers
              1..2402
              /organism="Homo sapiens"

```

```

/db_xref="taxon:9606"
/map="2p"
1..2402
/gene="SDC"
/mRNA
<1..2402
/gene="SDC"
/product="syndecan mRNA"
206..1138
/gene="SDC"
/note="syndecan"
/codon_start=1
/db_xref="GI:338634"
/db_xref="DB:600-126-375"
/translation="MRRALMLWLCALSLQALPQIVATNLPEDDQSGSDSDNF
SGSGALADLITLSQPTSTWMDTOLTAIPSPPEPTGLXTASTSTLPRGPRKG
EAVLPEVPEPGLTAEQATPRRETQPTTHQSTTTATTAQBPATSHPRDMQPG
HHETSPAPSGQDLHTPHTEDDGSPATERRAEDGASSQLPAAGSGEODFETSGE
NTAVVAPEPRNOSPVNDGATCASOGILDRREVLGVIAAGLVGLIFAVCLVGNLX
RMKKDGGSYSLPEPKANGAIYKPKOEERYA"
BASE COUNT      462 a 735 c 712 g 492 t 1 others
ORIGIN

```

```

Query Match      89.0%; Score 289.4; DB 9; Length 2402;
Best Local Similarity 99.0%; Pred. No. 7.7e-62;
Matches 301; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY      4  GAGGGCGGAGGGGCTGAGATCCTCTGACAGACACGCCCTCTGCTGAGCGCG 63
Db      2100 GAGGGCGGAGGGGCTGAGATCCTCTGACAGACACGCCCTCTGAGCGCG 2158

QY      64  TCTCCAGGGGCTGCTTCTCTGAAATTGACAGGGGTGCTTGGGAGAGTGGCTCT 123
Db      2159 TCTCCAGGGGCTGCTTCTCTGAAATTGACAGGGGTGCTTGGGAGAGTGGCTCT 2218

QY      124 GAGCGCTCCATCCAGGCGAGGTCTCCGTTAGCTCTGAGGCCACCTGGGCGCTG 183
Db      124 GAGCGCTCCATCCAGGCGAGGTCTCCGTTAGCTCTGAGGCCACCTGGGCGCTG 2218

QY      2219 GAGCGCTCCATCCAGGCGAGGTCTCCGTTAGCTCTGAGGCCACCTGGGCGCTG 2278
Db      2219 GAGCGCTCCATCCAGGCGAGGTCTCCGTTAGCTCTGAGGCCACCTGGGCGCTG 2278

QY      184 GGTGGATCAGGAATATTTTCCAAAGATGATAGTCTTTTGGCAAACTCTAC 243
Db      2279 GGTGGATCAGGAATATTTTCCAAAGATGATAGTCTTTTGGCAAACTCTAC 2338

QY      244 TTAATCAATGGGTTTTCTCTGACAGTAGATTTTCCAAATGTAAATACTTAATATA 303
Db      2339 TTAATCAATGGGTTTTCTCTGACAGTAGATTTTCCAAATGTAAATACTTAATATA 2398

QY      304 AAGT 307
Db      2399 AAGT 2402

```

```

RESULT 8
AX336413
LOCUS      AX336413      4797 bp      DNA      linear      PAT 09-JAN-2002
DEFINITION      Sequence 6922 from Patent WO0194629.
ACCESSION      AX336413
VERSION      AX336413.1 GI:18127132
KEYWORDS      human.
SOURCE      Homo sapiens
ORGANISM      Homo sapiens
REFERENCE      1
AUTHORS      Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Harrigan, S., Sopper, D.R. and Weaver, Z.
TITLE      Cancer gene identification and therapeutic screening using signature
JOURNAL      Patent: WO 0194629-A 6922 13-DEC-2001;
MEDLINE      Avalon Pharmaceuticals (US)
PUBMED      114797
COMMENT      Location/Qualifiers
FEATURES      1..4797
SOURCE      /organism="Homo sapiens"

```

```

BASE COUNT      957 a /db_xref="taxon:9606"      1349 c 1386 g 1099 t      6 others
ORIGIN
Query Match      87.9%; Score 285.8; DB 6; Length 4797;
Best Local Similarity 96.1%; Pred. No. 6,2e-61;
Matches 293; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 4 GAGGGGCGAGGGGCGCTGAGATCCTCTGACAGACCGCCGCTGCGCTGGGCGCG 63
DB 3765 GAGGGGCGAGGGGCGCTGAGATCCTCTGACAGACCGCCGCTGCGCTGGGCGCG 3824
QY 64 TCTCCAGGGGCTGCTTCTCTGGAATTGACAGGGGCTGCTTGGGAGAGCTGCTCT 123
DB 3825 TCTCCAGGGGCTGCTTCTCTGGAATTGACAGGGGCTGCTTGGGAGAGCTGCTCT 3884
QY 124 GAGCGCTTCATCCAGGCGAGTTCCTGCTTACCTGTCGCGCCACCTCGGCGCTG 183
DB 3885 GAGCGCTTCATCCAGGCGAGTTCCTGCTTACCTGTCGCGCCACCTCGGCGCTG 3944
QY 184 GCGTGAATCAGGAATATTTTCCAAAGATGATGCTTTGCTTGGCAAACTCTAC 243
DB 3945 GCGTGAATCAGGAATATTTTCCAAAGATGATGCTTTGCTTGGCAAACTCTAC 4004
QY 244 TTAATCCAAATGGGTTTCTCTGACAGTAGATTTTCCAAATGTAATTAATTATATA 303
DB 4005 TTAATCCAAATGGGTTTCTCTGACAGTAGATTTTCCAAATGTAATTAATTATATA 4064
QY 304 AAGTA 308
DB 4065 AAGTA 4069

RESULT 9
AX365741 4797 bp DNA linear PAT 15-FEB-2002
LOCUS AX365741
DEFINITION Sequence 134 from Patent WO0200174.
ACCESSION AX365741
VERSION AX365741.1 GI:18697290
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 Wang, T., Wang, A., Skeiky, Y. A., Li, S. X., Kalos, M. D., Henderson, R. A.,
McNeill, P. D., Fanger, N., Retter, M. W., Marnettakis, M., Fanger, G. R.,
Vedvick, T. S., Carter, D., Watanabe, Y., and Peckham, D. W.
Compositions and methods for the therapy and diagnosis of lung
cancer.
Patent: WO 0200174-A 134 03-JAN-2002;
JOURNAL CORIXA CORPORATION (US)
FEATURES
source 1..4797
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT      957 a 1349 c 1386 g 1099 t      6 others
ORIGIN
Query Match      87.9%; Score 285.8; DB 6; Length 4797;
Best Local Similarity 96.1%; Pred. No. 6,2e-61;
Matches 293; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```

```

DB 3885 GAGCGCTTCATCCAGGCGAGTTCCTGAGCTCCTGTCGCGCCACCTCGGCGCTG 3944
QY 184 GCGTGAATCAGGAATATTTTCCAAAGATGATGCTTTGCTTGGCAAACTCTAC 243
DB 3945 GCGTGAATCAGGAATATTTTCCAAAGATGATGCTTTGCTTGGCAAACTCTAC 4004
QY 244 TTAATCCAAATGGGTTTCTCTGACAGTAGATTTTCCAAATGTAATTAATTATATA 303
DB 4005 TTAATCCAAATGGGTTTCTCTGACAGTAGATTTTCCAAATGTAATTAATTATATA 4064
QY 304 AAGTA 308
DB 4065 AAGTA 4069

RESULT 10
AX411295 4797 bp DNA linear PAT 14-JUN-2002
LOCUS AX411295
DEFINITION Sequence 3942 from Patent WO0229103.
ACCESSION AX411295
VERSION AX411295.1 GI:21444000
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 Alvarez, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.
Gene expression profiles in liver cancer
Patent: WO 0229103-A 3942 11-APR-2002;
JOURNAL GENE LOGIC INC (US)
FEATURES
source 1..4797
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/Note="EMBL/Genbank Accession No. Z48199"

BASE COUNT      957 a 1349 c 1386 g 1099 t      6 others
ORIGIN
Query Match      87.9%; Score 285.8; DB 6; Length 4797;
Best Local Similarity 96.1%; Pred. No. 6,2e-61;
Matches 293; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 4 GAGGGGCGAGGGGCGCTGAGATCCTCTGACAGACCGCCGCTGCGCTGGGCGCG 63
DB 3765 GAGGGGCGAGGGGCGCTGAGATCCTCTGACAGACCGCCGCTGCGCTGGGCGCG 3824
QY 64 TCTCCAGGGGCTGCTTCTCTGGAATTGACAGGGGCTGCTTGGGAGAGCTGCTCT 123
DB 3825 TCTCCAGGGGCTGCTTCTCTGGAATTGACAGGGGCTGCTTGGGAGAGCTGCTCT 3884
QY 124 GAGCGCTTCATCCAGGCGAGTTCCTGCTTACCTGTCGCGCCACCTCGGCGCTG 183
DB 3885 GAGCGCTTCATCCAGGCGAGTTCCTGCTTACCTGTCGCGCCACCTCGGCGCTG 3944
QY 184 GCGTGAATCAGGAATATTTTCCAAAGATGATGCTTTGCTTGGCAAACTCTAC 243
DB 3945 GCGTGAATCAGGAATATTTTCCAAAGATGATGCTTTGCTTGGCAAACTCTAC 4004
QY 244 TTAATCCAAATGGGTTTCTCTGACAGTAGATTTTCCAAATGTAATTAATTATATA 303
DB 4005 TTAATCCAAATGGGTTTCTCTGACAGTAGATTTTCCAAATGTAATTAATTATATA 4064
QY 304 AAGTA 308
DB 4065 AAGTA 4069

RESULT 11
HSSYND1CN 4797 bp DNA linear PRI 09-FEB-1995
LOCUS HSSYND1CN
DEFINITION H.sapiens syndecan-1 gene (exons 2-5).
ACCESSION Z48199

```

FEATURES	source	location/Qualifiers
REFERENCE		1. 4797
AUTHORS		/organism="Homo sapiens"
TITLE		/db_xref="taxon:9606"
JOURNAL		/clone="cosmid 4.10"
AUTHORS		/issue_type="placenta"
REFERENCE		/clone_lib="Stratagene"
TITLE		/dev_slib="adult"
JOURNAL		join(177..258,1302..1780,2348..2483,2634..2803)
		/gene="syndecan-1"
		join(<177..258,1302..1780,2348..2483,2634..2803)
		/gene="syndecan-1"
		/codon_start=1
		/protein_id="CA88235.1"
		/db_xref="GI:666052"
		/db_xref="SWISS-PROT:P18827"
		/translation="QIVATNLPEDDQSGSDSNFSGSAGALODITLSDPTSTWKDTOLAIPTSPPEPGEATASTSLPAGBSGPKGSAVLPVEPGLTARQATPRPRETLOPTTHQASTTTATTAQEPATSHPRMOPGHETSPAGSOADLTHTEDGSEATERRARAGASQCPAABSGSDPTFSGENATVAAYEPRRQSGVDQATGASGLDKKEVLGSGVINGIVGLTAVCLVDFMLYMKKKDGSISLEPKQNGA YQRTKQEEFYA"
exon		177..258
		/gene="syndecan-1"
		/number=2
intron		258..1301
		/gene="syndecan-1"
		/number=2
exon		1302..1780
		/gene="syndecan-1"
		/number=3
intron		1781..2347
		/gene="syndecan-1"
		/number=3
exon		2348..2483
		/gene="syndecan-1"
		/number=4
intron		2484..2633
		/gene="syndecan-1"
		/number=4
exon		2634..4798
		/product="transmembrane and cytoplasmic domains"
		/number=5
3' UTR		2804..4798
polyA_signal		/note="first polyA signal"
polyA_signal		4733..4738
		/note="second polyA signal"
BASE COUNT	957 a 1349 c 1386 g 1039 t	6 others
ORIGIN		
Query Match	87.9%; Score 285.8; DB 9; Length 4797;	
Best Local Similarity	96.1%; Pred. No. 6.2e-61;	
Matches 293; Conservative	0; Mismatches 12; Indels 0; Gaps 0;	

	Db	3765	GAGGGGCGCAGAGGGGCCTGGAATCCCTCTCAAGGCTCAAGGCCCTCTCTGTGGGACGG	3824
	QY	64	TCTCCAGGGGCTCTTCTCTCTCGAAATTGACAGAGGGTGTCCTTGGGACAAGCTGGCTCT	123
	Db	3825	TCTCCAGGGGCTGTCTCTCTCTCGAAATTGACCAGAGGGTGTCCTTGGGACAAGCTGGCTCT	3884
	QY	124	GAGGGCTCCATCCAAAGGCGAGTTCCGTTAGCTCCCTGAGGCCACCGCTGGGCCCTG	183
	Db	3885	GAGGGCTCCATCCAAAGGCGAGTTCTCCGTTAGCTCTGAGGCCACCGCTGGGCCCTG	3944
	QY	184	GAGCTGGAATCAGAATAATTTTCCAAGAGTAGTAGCTTTTGCTTTTGGCAAAACTCTAC	243
	Db	3945	GAGCTGGAATCAGAATAATTTTCCAAGAGTAGTAGCTTTTGCTTTTGGCAAAACTCTAC	4004
	QY	244	TTAATCCATGGGTTTTTCTGTGACAGTAGAATTTTCCAATGTATPAACTTTAATATA	303
	Db	4005	TTAATCCATGGGTTTTTCTGTGACAGTAGAATTTTCCAATGTATPAACTTTAATATA	4064
	QY	304	AAGTA 308 	
	Db	4065	AAGTA 4069 	
	RESULT 12			
	AX454046			
	LOCUS	AX454046	425 bp	DNA linear PAT 06-JUL-2002
	DEFINITION	Sequence 22 from Patent WO0198539.		
	ACCESSION	AX454046		
	VERSION	AX454046.1 GI:21713684		
	KEYWORDS			
	SOURCE	'		
	ORGANISM	synthetic construct. synthetic construct artificial sequences.		
	REFERENCE	1		
	AUTHORS	Mitsunashi,M., Kambara,H., Matsunaga,H. and Kawamura,M.		
	TITLE	Gene markers for lung cancer		
	JOURNAL	Patent: WO 0198539-A-22 27-DEC-2001; Hitachi Chemical Co., Ltd. (JP); HITACHI CHEMICAL RESEARCH CENTER, INC. (US); Hitachi, Ltd. (JP) Location/Qualifiers 1..425 /organism="Synthetic construct" /db_xref="taxon:32630" /note="Syndecan 1 probe."		
	FEATURES			
	source			
	BASE COUNT	62 a 124 c 146 g 93 t		
	ORIGIN			
	Query Match	55.8%; Score 181.4; DB 6; Length 425;		
	Best Local Similarity	94.5%; Pred. No. 7.5e-35;		
	Matches 188; Conservative 0; Mismatches 11; Indels 0; Gaps 0			
	QY	4	GAGGGGCGAGGGGCTCGAGATCCTCTCGACAACAGCGCCCTCTGCTGTGGGCGCG	63
	Ddb	227	GAGGGGCGAGGGGCTCGAGATCCTCTCGACGAGCTACCGCTCTCTGTGGCGCG	286
	QY	64	TCTCCAGGGGCTCTCTCTCTCGAAATTGACAGAGGGTGTCCTTGGGAGAGCTGGCT	123
	Ddb	287	TCTCCAGGGGCTCTCTCTCTCGAAATTGACAGAGGGTGTCCTTGGGAGAGCTGGCT	346
	QY	124	GAGGGCTCCATCCAAAGGCGAGTTCTCCGTTAGCTCTGTGGCCCCAACCTGGGCGCTG	183
	Ddb	347	GAGGGCTCCATCCAAAGGCGAGTTCTCCGTTAGCTCTGTGGCCCCAACCTGGGCGCTG	406
	QY	184	GAGCTGGAATCAGAATAATTTTCCAAGAGTAGTAGCTTTTGCTTTTGGCAAAACTCTAC	202
	Ddb	407	GAGCTGGAATCAGAATAATTTTCCAAGAGTAGTAGCTTTTGCTTTTGGCAAAACTCTAC	425
	RESULT 13			
	AX368331		492 bp	DNA linear PAT 16-FEB-2001
	DEFINITION	Sequence 1041 from Patent WO0204514.		
	ACCESSION	AX368331		

```

VERSION      AX36831.1 GI:18856404
KEYWORDS
SOURCE       Homo sapiens
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1
AUTHORS      Wang, T., Watanabe, Y., Henderson, R.A., Johnson, J.C., Retter, M.W.,
              Manerakis, M., Carter, D., Fanger, G.R., Vedvyck, T.S., Bangur, C.S.,
              Mcnabb, A., Fanger, N., Switzer, A., McNeill, P.D. and Clapper, J.D.
TITLE        Compositions and methods for the therapy and diagnosis of lung
              cancer
JOURNAL      Patent: WO 0204514-A 1041 17-JAN-2002;
              CORNIXA CORPORATION (US)
FEATURES
  source      Location/Qualifiers
              1..492
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
BASE COUNT   153 a 126 c 110 g 102 t 1 others
ORIGIN
Query Match 44.4%; Score 144.4; DB 6; Length 492;
Best Local Similarity 98.6%; Pred. No. 1.4e-25;
Matches 145; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 162 TGTGGCCCACTGGGCGCTGGGAGATCAGGAATATTTTCCAAAGATGATGCT 221
Db 492 TGTGGCCCACTGGGCGCTGGGAGATCAGGAATATTTTCCAAAGATGATGCT 433

Qy 222 TTGCTTTGGCAAACTCTACTTAATCCATGAGTTTCTCTGACATAGATTTCC 281
Db 432 TTGCTTTGGCAAACTCTACTTAATCCATGAGTTTCTCTGACATAGATTTCC 373

Qy 282 AAATGTAATAAATTAAATATAAGTA 308
Db 372 AAATGTAATAAATTAAATATAAGTA 346

RESULT 14
AL672026/c 200548 bp DNA linear HTG 17-AUG-2002
LOCUS      Mus musculus chromosome X clone RP23-403011, *** SEQUENCING IN
DEFINITION PROGRESS ***, in ordered pieces.
ACCESSION  AL672026 GI:22204360
VERSION     AL672026.9
KEYWORDS    HTG; HTGS_PHASE2; HTGS_ACTIVEPIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      house mouse.
ORGANISM    Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 200548)
REFERENCE   1
AUTHORS     Tromans, A.
TITLE       Direct Submission
JOURNAL     Submitted (16-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,
              Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
              humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
              On Aug 11, 2002 this sequence version replaced gi:21955346.
              ----- Genome Center
              Center: Wellcome Trust Sanger Institute
              Center code: SC
              Web site: http://www.sanger.ac.uk
              Contact: humquerry@sanger.ac.uk
              ----- Project Information
              Center project name: BM403011
              ----- Summary Statistics
              Assembly program: XGAP4; version 4.5
              Chemistry: Dye-terminator; 5% of reads
              Chemistry: Dye-terminator Big Dye; 94% of reads
              Consensus quality: 200513 bases at least Q40
              Consensus quality: 200548 bases at least Q30
              Consensus quality: 200548 bases at least Q20
              Insert size: 200548; sum-of-contrigs
              Insert size: 172347; 7.0% error; agarose-fp

```

```

Quality coverage: 8.27x in Q20 bases; sum-of-contrigs Quality
coverage: 9.63x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
Location/Qualifiers
  source      1..200548
              /organism="Mus musculus"
              /db_xref="taxon:10090"
              /chromosome="X"
              /clone="RP23-403011"
              /clone_lib="RPC1-23"
              misc_feature 1..200548
              /note="assembly_fragment:03152"
BASE COUNT   56146 a 44414 c 44985 g 55003 t
ORIGIN
Query Match 40.0%; Score 130; DB 2; Length 200548;
Best Local Similarity 76.1%; Pred. No. 6.4e-22;
Matches 216; Conservative 0; Mismatches 55; Indels 13; Gaps 4;

Qy 38 CCAAGCCCTCTGCTGCTGAGCGCGCTCCAGAGGCTCTTCTCTGGAATTGACGA 97
Db 17768 CCAAGGCTCTCCACTTTGGTACCATCTCTAGTCAC-CTTTCTCTCTGGAATTGACGA 177710

Qy 98 GGGGTCTCTGGGCAAGAGCTGCTGAGCGCTCCATTCAGAGCCAGGCTTCCGTTG 157
Db 177709 GACACATCTTAATGATGCTGAGCAGTG-GTTCCTCATCAAGAACCAAGTTCACTTCG 177651

Qy 158 CTCCTGTGCCCCCACTGGGCGCTGGGAGATCAGGAATATTTTCCAAAGATGATA 217
Db 177650 CTCCTGTGCCCCC-----CCCAAGCTGAGATCAAGAAATTTTCCCAAGATGATA 177597

Qy 218 GTCTTTGCTTTGGCAAACTCTACTTAATCCAAATGGGTTTCTCTGACATAGATT 277
Db 177596 GTCTTTGCTTTGGCAAAAGCTACTTAATCCAAATGGGTTTCTCTGACATAGATT 177542

Qy 278 TTCCAAATGTAATAAATTAAATATAAGTAAGTAAGTAAGTAAGTAAGTAAGTA 321
Db 177541 TTCCAAATGTAATAAATTAAATATAAGTAAGTAAGTAAGTAAGTAAGTA 177498

RESULT 15
AC002406 194985 bp DNA linear ROD 21-MAR-1998
LOCUS      Mouse chromosome X BAC B178A13 (Research Genetics mouse BAC
DEFINITION library) complete sequence.
ACCESSION  AC002406
VERSION     AC002406.1 GI:2981248
KEYWORDS    HTG.
SOURCE      Mus musculus.
ORGANISM    Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 194985)
REFERENCE   1
AUTHORS     Muzny, D., Arenson, A.D., Brundage, E., Carvelli, K., Chen, E., Chen, J.,
              Di, W., Ding, Y., Dugan, S., Durbin, J., Forcum, J., Ganesh, R.,
              Garcia, C., Goodman, M., Gorrell, J.H., Haywood, M., Jackson, L.,
              Jin, S., Kampal, R., Karpach, S., Leal, B., Li, Y., Liu, N., Logan, O.,
              Lu, Y., Ly, T., Martinez, C., Oswal, G., Perez, L., Rashid, N.D.,
              Rowland, K., Savage, L., Scherer, S.S., Shen, H., Timms, K.M., Todd, J.,
              Vo, Q., Worley, K.C., Yu, W., Chinalut, C., Nelson, D. and Gibbs, R.A.
TITLE       Direct Submission
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 194985)
AUTHORS     Chin, M.W.
TITLE       Direct Submission
JOURNAL     Submitted (11-AUG-1997) Molecular and Human Genetics, Baylor
              College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3 (bases 1 to 194985)
AUTHORS     Worley, K.C.

```

TITLE

Direct Submission
Submitted (21-MAR-1998) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Mar 23, 1998 this sequence version replaced gi:2909685.

Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality does not meet this standard, it will be indicated in the annotation.

The repeat regions shown were identified using RepeatMasker by Adrian Smit.

Sequence similarities were identified using Powerblast by Jinghui Zhang.

Exon/Intron boundaries of identified genes were chosen if there were canonical splice junctions that maintained sequence continuity across the splice junctions.

FEATURES

```

source
    1..194985
    /organism="Mus musculus"
    /db_xref="taxon:10090"
    /chromosome="X"
    /clone="B178A13"
    /clone_lib="Research Genetics mouse BAC library"
    226..352
    repeat_region
        /rpt_family="MSTD"
        562..599
        repeat_region
            /rpt_family="(TA)n"
            complement(863..1868)
            /rpt_family="L1_MM"
        repeat_region
            complement(2661..3921)
            /rpt_family="B4A"
        repeat_region
            complement(5160..5372)
            /rpt_family="B3"
        repeat_region
            complement(5456..5705)
            /rpt_family="L1MD3"
        repeat_region
            complement(5766..5807)
            /rpt_family="(TAAA)n"
        repeat_region
            complement(5900..6083)
            /rpt_family="L1MD3"
        repeat_region
            complement(6433..6546)
            /rpt_family="(TAAA)n"
        repeat_region
            complement(6547..6575)
            /rpt_family="AT_rich"
        repeat_region
            complement(6576..6776)
            /rpt_family="L1MD2"
        repeat_region
            complement(6808..6947)
            /rpt_family="B1-F"
        repeat_region
            complement(7019..7348)
            /rpt_family="L1MD2"
        repeat_region
            7396..7541
            /rpt_family="B1_MM"
        repeat_region
            7542..7587
            /rpt_family="(CAA)n"
        repeat_region
            complement(7915..8149)
            /rpt_family="L1M4"
        repeat_region
            8166..8512
            /rpt_family="L1M4"
        repeat_region
            complement(8730..9162)
            /rpt_family="L1_MM"
        repeat_region
            9149..9265
            /rpt_family="L1_MM"
        repeat_region
            complement(9454..9606)
            /rpt_family="RLTR10"
        repeat_region
            complement(9734..10166)
            /rpt_family="L1_MM"
        repeat_region
            11635..11661
            /rpt_family="POLY_A"
        repeat_region
            complement(11830..11966)
            /rpt_family="RLTR10"

repeat_region
    11967..12203
    /rpt_family="L1_MM"
repeat_region
    12575..12708
    /rpt_family="B1-F"
repeat_region
    12711..12772
    /rpt_family="(CAAAA)n"
    complement(13570..13598)
    /rpt_family="AT_rich"
repeat_region
    14287..14359
    /rpt_family="MIR"
repeat_region
    14418..15428
    /rpt_family="L1"
repeat_region
    15490..15596
    /rpt_family="L1"
repeat_region
    15828..16199
    /rpt_family="Lx4"
    complement(16753..16830)
    /rpt_family="MTE"
repeat_region
    17198..17275
    /rpt_family="MLT1A1"
repeat_region
    17311..17443
    /rpt_family="MLT1A2"
    complement(19559..19879)
    /rpt_family="MTE"
repeat_region
    complement(19895..19963)
    /rpt_family="MTE"
    19964..26215
    /rpt_family="L1F"
    complement(26283..26511)
    /rpt_family="L1"
repeat_region
    26468..27422
    /rpt_family="L1_MM"
    complement(27624..27755)
    /rpt_family="B1-F"
    27759..28051
    /rpt_family="MTC"
    28280..28490
    /rpt_family="L1"
    28491..33172
    /rpt_family="L1_MM"
    33178..33672
    /note="Region: Similar to U29402|NMU29402 Mus musculus
    acidic ribosomal phosphoprotein P1 mRNA, complete cds"
repeat_region
    33682..35079
    /rpt_family="L1_MM"
    37207..37292
    /rpt_family="RSINE1"
    complement(37310..37338)
    /rpt_family="(GA)n"
    37354..37474
    /rpt_family="(CAT)n"
    complement(37574..37728)
    /rpt_family="RSINE1"
    complement(38061..38179)
    /rpt_family="(TA)n"
    39320..39492
    /rpt_family="B4"
repeat_region
    39603..39661
    /rpt_family="(CA)n"
    complement(40026..40452)
    /rpt_family="Lx2B"
    40580..40624
    /rpt_family="(CA)n"
    complement(40951..41040)
    /rpt_family="MLT2B"
    41143..41538
    /rpt_family="MTA"
    complement(42185..42262)
    /rpt_family="(CAG)n"
    complement(43208..43318)
    /rpt_family="(CAGA)n"
    complement(43486..43974)
    /rpt_family="L1"

```

```
repeat_region      43954. .44894  
/rpt_family="L1_MM"  
repeat_region      45112. .45198  
/rpt_family="(CAA)n"  
repeat_region      complement(45459. .45479)  
/rpt_family="AT_rich"  
repeat_region      46721. .46836  
/rpt_family="(GAA)n"  
repeat_region      46914. .47150  
/rpt_family="MTD"  
repeat_region      47213. .47296  
/rpt_family="PB1D10"  
repeat_region      50241. .50324  
/rpt_family="MTC"  
repeat_region      50328. .51555  
/rpt_family="L1_MM"  
repeat_region      51568. .52378  
/rpt_family="L1_MM"  
repeat_region      52399. .52434  
/rpt_family="POLY_A"  
repeat_region      52446. .52503  
/rpt_family="L1_MM"  
repeat_region      52520. .52838  
/rpt_family="MTC"  
repeat_region      53200. .53245  
/rpt_family="(AAAA)n"  
repeat_region      complement(53474. .53534)
```

Query Match 39.9%; Score 129.8; DB 10; Length 194985;

Best Local Similarity 75.6%; Pred. No. 7.1e-22;
Matches 217; Conservative 0; Mismatches 57; Indels 13; Gaps 4;

```
QY 38 CCACGCCCGTCTGCGCTGGGGCGCGCTCCAGGGGCTGCTCTCTCTGGAATTGACGA 97  
DB 148547 CCAAGCGTCTCCACCTTGACATCTAGTAC -CCTTCTCCCGAAGTTGACAA 148605  
QY 98 GGGGTGCTGGGCAAGTGGCTGTGAGCGCCTCCATCCAGGCCAGTTCTCCGTTAG 157  
DB 148606 GACACATCTTGAATGATGGCTGCACTG -GTTCCTCATCAAGAACCAAGTTCACTTCAG 148664  
QY 158 CTCTGTGGGCCCACTGGGCGCTGGGCTGGAATCAGGATATTTTCCAAAGAGTGATA 217  
DB 148665 CTCTGTGGGCCCGCCT-----CGAGGCTGAGTCAAGAAATGTTCCCAAGAGTGAGA 148718  
QY 218 GTCTTTTGCTTTTGGCAAACTTACTTAATCAATGGGTTTCTCTGTACAGTACATT 277  
DB 148719 GTCTTTTGCTTTTGGCAAACTTACTTAATCAATGGGTTTCTCTGTACAGTACATT 148773  
QY 278 TTCCAAATGTAATTAATTTAATATAAGTAAAAAAGAGAGAGAGAGAGAGAGAGAG 324  
DB 148774 TTGCAGATGTAATTAATTTAATATAAGTAAAAAAGAGAGAGAGAGAGAGAGAG 148820
```

Search completed: January 10, 2003, 02:45:16
Job time : 877.023 secs


```

XX 13-JUN-1995; 95NO-FI00344.
PF 13-JUN-1994; 94US-0258862.
PR 13-JUN-1994; 94US-0258862.
XX
XX (JALK/) JALKANEN M.
PA (MALI/) MALI M.
XX
XX Jalkanen M., Mali M;
PI
XX WPI: 1996-049416/05.
DR P-PSDB; AAR87000.
XX
XX New syndecan ectodomain compsn. to reduce tumour growth - for
PT delivery to extracellular environment for suppressing tumour growth
PT in malignant and non-malignant tumours
XX
XX Disclosure: Fig 1; 49pp; English.
XX
XX A human DNA (AAT08125) codes for syndecan-1 (AAR87000), a cell surface
CC proteoglycan. The ectodomain of syndecan-1 (amino acids 18-251)
CC is capable of restoring a more differentiated morphology to tumour
CC cells and to suppress the growth of malignant cells. The DNA
CC can be used for the prodn. of recombinant ectodomain, pref. using
CC host cells of a similar cell type as that of the tumour. The
CC mouse syndecan-1 gene (AAT08126) has also been identified.
XX
SQ Sequence 2430 BP; 490 A; 736 C; 711 G; 492 T; 1 other;

Query Match 94.6%; Score 307.4; DB 17; Length 2430;
Best Local Similarity 99.1%; Pred. No. 4.5e-68;
Matches 319; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 4 GAGGGGGGAGGGGGCTGGAGATCTCTGACAGACAGCCCGCTGCTGGGCGCG 63
DB 2100 GAGGGGGGAGGGGGCTGGAGATCTCTGACAGACAGCCCGCTGCTGGGCGCG 2158
QY 64 TCTCCAGGGGGCTGCTCTCTCTGGAATTGACAGGGGGTGTCTTGGGACAGCTGCTCT 123
DB 2159 TCTCCAGGGGGCTGCTCTCTCTGGAATTGACAGGGGGTGTCTTGGGACAGCTGCTCT 2218
QY 124 GAGCGCTTCATCCAGGCGCAGGTTCTCCGTTAGCTCTGCGGCCCAACCTGGGCCCTG 183
DB 2219 GAGCGCTTCATCCAGGCGCAGGTTCTCCGTTAGCTCTGCGGCCCAACCTGGGCCCTG 2278
QY 184 GCGTGAATCAGGAATATTTTCCAAAGAGTAGTCTTTGGCAAAACTCTAC 243
DB 2279 GCGTGAATCAGGAATATTTTCCAAAGAGTAGTCTTTGGCAAAACTCTAC 2338
QY 244 TTAATCCAATGGGTTTTTCTGTACAGTAGATTTTCCAAATGTAATPAACTTAATATA 303
DB 2339 TTAATCCAATGGGTTTTTCTGTGTACAGTAGATTTTCCAAATGTAATPAACTTAATATA 2398
QY 304 AAGTAAAAAAAAAAAAAAAAA 325
DB 2399 AAGTAAAAAAAAAAAAAAAAA 2420

RESULT 2
ID AAV81282 standard; DNA; 2430 BP.
XX
XX AAV81282;
XX
XX 11-MAR-1999 (first entry)
DB Human syndecan-1 encoding DNA.
XX
XX S ndecan-1; tumour; ectodomain; epithelial; mesenchymal; breast; human;
KM endometrial tumour; prostatic carcinoma; oestrogenic; androgenic; steroid;
KM glioma; myeloma; carcinoma; sarcoma; lymphoma; adenoma; ds.
XX
XX Homo sapiens.

```

```

XX Key Location/Qualifiers
FH 206..1138
FT CDS
FT
FT /*tag= a
FT /product= "human syndecan-1"
FT polyA_signal
FT 2383..2388
FT /*tag= b
FT /note= "q"
XX
XX US585193-A.
XX
XX 22-DEC-1998.
XX
XX 07-JUN-1995; 95US-0488139.
PF 07-JUN-1995; 95US-0488139.
XX
XX 07-JUN-1995; 95US-0488139.
PR 13-JUN-1994; 94US-0258862.
XX
XX (BIOT-) BIOTIE THERAPIES LTD.
PA
XX
XX Jalkanen M., Mali M;
PI
XX
XX WPI: 1999-104635/09.
DR P-PSDB; AAW95197.
XX
XX Reducing growth of tumour cells - with ectodomain of syndecan
PT applied to the extracellular environment to induce a more
PT differentiated phenotype, particularly for hormone-dependent breast,
PT endometrial or prostatic cancers
XX
XX Claim 7; Fig 1A-B; 48pp; English.
XX
XX This DNA encodes a human syndecan-1 protein. This is used in the method
CC of the invention of reducing the growth of tumour cells. The method
CC comprises supplying a syndecan ectodomain to the environment around the
CC cells causing them to develop a more differentiated phenotype. The method
CC is used to suppress tumour cells of epithelial, mesenchymal, pre-B or
CC plasma cell origin, especially breast, endometrial or prostatic tumours,
CC and particularly those responsive to an oestrogenic or androgenic
CC steroid. More generally it can be used to treat malignant or non-
CC malignant tumours, particularly those characterised by loss of syndecan,
CC e.g. gliomas, myelomas, carcinomas, sarcomas, lymphomas and adenomas.
XX
SQ Sequence 2430 BP; 490 A; 735 C; 712 G; 492 T; 1 other;

Query Match 94.6%; Score 307.4; DB 20; Length 2430;
Best Local Similarity 99.1%; Pred. No. 4.5e-68;
Matches 319; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 4 GAGGGGGGAGGGGGCTGGAGATCTCTGACAGACAGCCCGCTGCTGGGCGCG 63
DB 2100 GAGGGGGGAGGGGGCTGGAGATCTCTGACAGACAGCCCGCTGCTGGGCGCG 2158
QY 64 TCTCCAGGGGGCTGCTCTCTCTGGAATTGACAGGGGGTGTCTTGGGACAGCTGCTCT 123
DB 2159 TCTCCAGGGGGCTGCTCTCTCTGGAATTGACAGGGGGTGTCTTGGGACAGCTGCTCT 2218
QY 124 GAGCGCTTCATCCAGGCGCAGGTTCTCCGTTAGCTCTGCGGCCCAACCTGGGCCCTG 183
DB 2219 GAGCGCTTCATCCAGGCGCAGGTTCTCCGTTAGCTCTGCGGCCCAACCTGGGCCCTG 2278
QY 184 GCGTGAATCAGGAATATTTTCCAAAGAGTAGTCTTTGGCAAAACTCTAC 243
DB 2279 GCGTGAATCAGGAATATTTTCCAAAGAGTAGTCTTTGGCAAAACTCTAC 2338
QY 244 TTAATCCAATGGGTTTTTCTGTACAGTAGATTTTCCAAATGTAATPAACTTAATATA 303
DB 2339 TTAATCCAATGGGTTTTTCTGTGTACAGTAGATTTTCCAAATGTAATPAACTTAATATA 2398
QY 304 AAGTAAAAAAAAAAAAAAAAA 325
DB 2399 AAGTAAAAAAAAAAAAAAAAA 2420

```


XX	RESULT 3
XX ID	AAFI17989
XX AC	AAFI17989 standard; DNA; 1763 BP.
XX	AAFI17989;
XX DT	14-MAR--2001 (first entry)
XX DE	Lung cancer associated polynucleotide sequence SEQ ID 8.
XX	Human; lung cancer associated protein; neuroprotective; cytosolic;
KW	cardioactive; immunomodulatory; muscular active; vulnary;
KW	gastrointestinal; nephrotoxic; antineutrophilic; gynecological;
KW	antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
KW	proliferative disorder; wound healing; infectious disease; ds.
XX	
OS	Homo sapiens.
XX	
XX	WO2000055180-A2.
XX PD	21-SEP-2000.
XX PF	08-MAR-2000; 2000MO-US05918.
XX PR	12-MAR-1999; 99US-0124270.
XX PA	(HUMA-) HUMAN GENOME SCI INC.
XX PA	(ROSE/) ROSEN C A.
XX PI	Ruben SM;
XX PS	WPI, 2000-587514/55.
DR	P-PSDB; AAB58113.
XX	
PT	Lung cancer associated gene sequences, referred to as lung cancer
PT	antigens, useful for treatment, prevention, and diagnosis of disorders
PT	such as lung cancer -
XX	
XX	Claim 1; Page 495; 1425pp; English.
XX	Polynucleotide sequences AAFI17982 - AAFI18424 encode human lung cancer
CC	associated proteins represented in AAB58106 - AAB58548. Lung cancer
CC	associated proteins and polynucleotide sequences, their agonists, and
CC	antagonists may have neuroprotective, cytosolic; cardioactive; and
CC	immunomodulatory; muscular active general; vulnary; gastrointestinal
CC	general; nephrotoxic; antineutrophilic; gynecological; or antibacterial
CC	activity. The invention also includes antibodies specific for the
CC	protein or polynucleotide sequences. The lung cancer associated
CC	polynucleotide sequences may be used for detection of lung cancer,
CC	chromosome identification, as chromosome markers, and for numerous other
CC	diagnostic or research purposes. The proteins may be used to treat
CC	disorders such as neural, immune, muscular, reproductive,
CC	gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC	disorders. The proteins may also be used in the treatment of wounds and
CC	infectious diseases. Polynucleotide sequences AAFI18425 - AAFI18433 and
CC	peptide AAB58549 are used in the course of the invention for the
CC	identification and characterisation of the polynucleotide and protein
CC	sequences.
XX	
XX	Sequence 1763 BP; 344 A; 450 C; 499 G; 470 T; 0 other;
XX	
QY	Query Match 89.7%; Score 291.4; DB 21; Length 1763;
Db	Best Local Similarity 99.3%; Pred. NO. 4.5e-64;
	Matches 303; Conservative 0; Mismatches 1; Indels 1; Gaps 1
QY	4 GAGGGGCGGAGGGGCTGGAATCTCTCGAGACCAAGCCGCTGCTGTGGCGCGG 63
Db	752 GAGGGGCGGAGGGGCTGGAATCTCTCGAGACCAAGCCGCTGCTGTGGCGCGG 811
QY	64 TCTTCAGGGGCTGCTTCTCTCTGGAATTCAGAGGGGTGTCTTGGCGAGACTGGCTCT 123
Db	812 TCTTCAGGGGCTGCTTCTCTCTGGAATTCAGAGGGGTGTCTTGGCGAGACTGGCTCT 871

QY	124	GAGGCGCCATCCATCCAAAGGCGAGGCTTCCTCCGTTAGCTCTGTGTGGCCCAACCCCTGGGCGCTTG	183
Db	872	GAGCGCTTCATCCAAAGCGCAGAGTTCCTCGTTAGCTCTGT- GCCCAACCCCTGGGCGCTTG	930
QY	184	GGCTTGATCAGGAATATTTTCCAAAGATGATAGTCTTTTGCTTTGGCAAACTCTAC	243
Db	931	GGCTTGATCAGGAATATTTTCCAAAGATGATAGTCTTTTGCTTTGGCAAACTCTAC	990
QY	244	TTAATCCATGGGTTTTCTCTGTACAGTACATTTTCCAAATGTATTAATTAATATA	303
Db	991	TTAATCCATGGGTTTTCTCTGTACAGTACATTTTCCAAATGTATTAATTAATATA	1050
QY	304	AAATA 308	
Db	1051	AAATA 1055	
RESULT 4			
AA001308	AA001308	standard; cDNA; 320 BP.	
XX	AC	AA001308;	
XX	DT	16-JUL-2001 (first entry)	
XX	DE	Bladder cancer-associated sequence, TCC75B3.	
XX	KW	Bladder; cancer; transitional cell carcinoma; diagnostic; TCC75B3; ss.	
XX	OS	Homo sapiens.	
XX	Key	Location/Qualifiers	
FT	FT	polyA_site	311
XX	PN	/*tag= a	
XX	XX	W0200122864-A2.	
XX	PD	05-APR-2001.	
XX	PF	27-SEP-2000; 2000MO-US41005.	
XX	XX	27-SEP-1999; 99US-0156153.	
PA	(OUAR-)	OUAR BIOTECH INC.	
FA	(KOHK-)	KOHK K I.	
XX	PI	Feinstein E, Mor O;	
XX	DR	WPI; 2001-258076/26.	
XX	XX		
PS	Claim 4; Page 56; 64p; English.		
CC	XX	The sequence represents the coding sequence of bladder cancer-associated sequence, transitional cell carcinoma clone, TCC75B3. The sequence is upregulated in bladder cancer and its expression is indicative of bladder cancer. The sequence can be used as a marker, and can be used for diagnosing bladder cancer. Antibodies, ribozymes, antisense oligonucleotides or a dominant negative peptide directed against the sequence are useful for regulating bladder cancer-associated pathologies in a patient.	
SQ	Sequence 320 BP; 73 A; 82 C; 81 G; 84 T; 0 other;		
QY	Query Match	88.4%; Score 287.2; DB 22; Length 320;	
	Best Local Similarity	97.8%; Pred. No. 2.9e-63;	
	Matches 313; Conservative	0; Mismatches 3; Indels 4; Gaps 2	

Db 1 AAAGAGGCGCGAGGCGCTGAGATCCTCTGACAGACACGCCCTGCTGCGG 60
 Qy 61 CCGTCTCCAGGCGCTGCTCTCTGGAATGACGAGGGGTGTCTTGGGACAGCTGAC 120
 Db 61 CCGTCTCCAGGCGCTGCTCTCTGGAATGACGAGGGGTGTCTTGGGACAGCTGAC 120
 Qy 121 TCTGAGC-GCTTCATCCAGGCGAGTTCCTCGTTAGCTCTTGCGCCCACTGCG 178
 Db 121 TCTGAGCGCGCTTCATCCAGGCGAGTTCCTCGTTAGCTCTTGCGCCCACTGCG 180
 Qy 179 CCGTGGCTGGAATCGAATATTTTCCAAAGAGTATGCTTTT-GCTTTTGGCAAA 236
 Db 181 CCGTGGCTGGAATCGAATATTTTCCAAAGAGTATGCTTTTGGCTTTGGCAAA 240
 Qy 237 ACTCTACTTAATCCATGAGGTTTCTCTGTACAGTAGATTTTCCAAATGTAATACTT 296
 Db 241 ACTCTACTTAATCCATGAGGTTTCTCTGTACAGTAGATTTTCCAAATGTAATACTT 300
 Qy 297 TTAATATAAGTAAAAAAA 316
 Db 301 TTAATATAAGTAAAAAAA 320

RESULT 5

AA65870
 ID AAC65870 standard, cDNA, 4797 BP.

AC AAC65870;

DT 21-FEB-2001 (first entry)

DE Human lung cancer-associated cDNA for contig 31.

KM Lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic;

KW vaccine; detection; ss.

OS Homo sapiens.

PN W0200061612-A2.

PD 19-OCT-2000.

PF 03-APR-2000; 2000MO-US08896.

PR 02-APR-1999; 99US-0285479.

PR 17-DEC-1999; 99US-0466396.

PR 30-DEC-1999; 99US-0476496.

PR 10-JAN-2000; 2000US-0480884.

PR 22-FEB-2000; 2000US-0510376.

PA (CORI-) CORIXA CORP.

PI Wang T, Fan J;

DR WPI; 2000-628399/60.

PT Isolated polypeptide comprising an immunogenic portion of a lung tumor

PT protein is used for detecting and monitoring progression of lung cancer

PT in a patient -

PS Claim 25a; Page 155-156; 261pp; English.

CC This invention describes a novel isolated polypeptide (1) which
 CC comprising an immunogenic portion of a lung tumor protein or variant (P2)
 CC which have cytostatic activity. The polypeptides and polynucleotides are
 CC used in compositions and vaccines to inhibit the development of cancer,
 CC especially lung cancer, in a patient. Methods described in the invention
 CC can be used to monitor the progression of a cancer by carrying out the
 CC detection at subsequent time points and comparing the results from the
 CC different time points. CD4+ and/or CD8+ T-Cells isolated from a patient
 CC are treated with P2, polynucleotides encoding P2 or antigen presenting
 CC cells expressing P2 and then administered to the patient to inhibit

CC development of cancer.

XX Sequence 4797 BP; 957 A; 1349 C; 1386 G; 1099 T; 6 other;

SQ Query Match 87.9%; Score 285.8; DB 21; Length 4797;

Best Local Similarity 96.1%; Pred. No. 1.6e-62;

Matches 293; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 4 GAGGCGGAGGCGCGTGGAGATCTCTGACAGACACGCCCTGCTGCGCGCG 63
 Db 3765 GAGGCGGAGGCGCGTGGAGATCTCTGACAGAGTCAAGCGCGCTCTCTGCGCGCG 3824
 Qy 64 TCTCCAGGCGCGTCTCTCTCTGGAATGACAGGCGGTCTTTGGGCAAGCTGCTCT 123
 Db 3825 TCTCCAGGCGCGTCTCTCTCTGGAATGACAGGCGGTCTTTGGGCAAGCTGCTCT 3884
 Qy 124 GAGCGCTTCATCCAGGCGAGGTTCTCGTTAGCTCTTGCGCCCACTGCGCG 183
 Db 3885 GAGCGCTTCATCCAGGCGAGGTTCTCGTTAGCTCTTGCGCCCACTGCGCGCG 3944
 Qy 184 GCGTGGATCGAAGATATTTTCCAAAGAGTATGCTTTTGGCAAACTCTAC 243
 Db 3945 GCGTGGATCGAAGATATTTTCCAAAGAGTATGCTTTTGGCAAACTCTAC 4004
 Qy 244 TTAATCCATGAGGTTTCTCTGTACAGTAGATTTTCCAAATGTAATACTTAAATA 303
 Db 4005 TTAATCCATGAGGTTTCTCTGTACAGTAGATTTTCCAAATGTAATACTTAAATA 4064
 Qy 304 AAGTA 308
 Db 4065 AAGTA 4069

RESULT 6

ABN97444
 ID ABN97444 standard; DNA; 4797 BP.

AC ABN97444;

DT 13-AUG-2002 (first entry)

DE Gene #3942 used to diagnose liver cancer.

KM Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;

KW metastatic liver tumour; cytostatic; expression profile; disease state;

OS disease progression; drug toxicity; drug efficacy; drug metabolism.

OS Homo sapiens.

PN W0200229103-A2.

PD 11-APR-2002.

PF 02-OCT-2001; 2001MO-US30589.

PF 02-OCT-2000; 2000US-237054P.

PA (GENE-) GENE LOGIC INC.

PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;

DR WPI; 2002-426119/45.

PT Diagnosing and detecting the progression of liver cancer,

PT hepatocellular carcinoma or metastatic liver tumor in a patient,

PT involves detecting the level of expression of two or more genes in a

PT liver tissue sample -

PS Claim 1; SEQ ID NO 3942; 298bp; English.

CC The invention relates to a novel method for diagnosing and detecting the

CC progression of liver cancer, hepatocellular carcinoma or metastatic liver

CC tumour in a patient, and differentiating metastatic liver cancer from

CC hepatocellular carcinoma in a patient, involving detecting the level of
 CC expression of two or more genes represented in ABN93503-ABN97455 in a
 CC tissue sample. The method of the invention has hepatotropic, and
 CC cytostatic activity. The method is useful for diagnosing and detecting
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic
 CC liver carcinoma in a patient. The method is useful for identifying
 CC expression profiles which serve as useful diagnostic markers as well as
 CC markers that can be used to monitor disease states, disease progression,
 CC drug toxicity, drug efficacy and drug metabolism.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_poc_sequences.
 XX
 SQ Sequence 4797 BP; 957 A; 1349 C; 1366 G; 1099 T; 6 other;
 Query Match 87.3%; Score 285.8; DB 24; Length 4797;
 Best Local Similarity 96.1%; Pred. No. 1.6e-62;
 Matches 293; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
 QY 4 GAAGGGGAGGGGGCTGGAGATCCCTGACAGACCGCCCTGCTGGGGCGG 63
 Db 3765 GAAGGGGAGGGGGCTGGAGATCCCTGACAGACCGCCCTGCTGGGGCGG 3824
 QY 64 TCTCCAGGGGCTGCTCTCTCTCTGGAATTGACAGGGGCTGCTGGGACAGCTGGCTCT 123
 Db 3825 TCTCCAGGGGCTGCTCTCTCTCTGGAATTGACAGGGGCTGCTGGGACAGCTGGCTCT 3884
 QY 124 GACGGCTTCATCCAGGCGCAGGTTCTCGTAGCTCTGTGGCCGACCCCTGGGCGCTG 183
 Db 3885 GACGGCTTCATCCAGGCGCAGGTTCTCGTAGCTCTGTGGCCGACCCCTGGGCGCTG 3944
 QY 184 GGCTGGAATCAGGAATATTTTCCAAAGAGTAGCTCTTTGGCTTTGGCAAACTGAC 243
 Db 3945 GGCTGGAATCAGGAATATTTTCCAAAGAGTAGCTCTTTGGCTTTGGCAAACTGAC 4004
 QY 244 TTATTCATAGGCTTTTCTCTGTCGATGATGATTTTCCAAATGATATTAATTATATA 303
 Db 4005 TTATTCATAGGCTTTTCTCTGTCGATGATGATTTTCCAAATGATATTAATTATATA 4064
 QY 304 AAGTA 308
 Db 4065 AAGTA 4069
 AC ABL68585;
 XX
 DT 15-MAY-2002 (first entry)
 XX
 DE Kidney cancer related gene sequence SEQ ID NO:6922.
 XX
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytoskeletal; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
 KW gene; de.
 XX
 OS Homo sapiens.
 XX
 FN WO200194629-A2.
 PD 13-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US10838.
 XX
 PR 05-JUN-2000; 2000US-209473P.
 PR 05-JUN-2000; 2000US-209531P.
 PR 18-SEP-2000; 2000US-231133P.
 PR 18-SEP-2000; 2000US-233617P.
 PR 20-SEP-2000; 2000US-234009P.
 PR 20-SEP-2000; 2000US-234034P.

PR 20-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234509P.
 PR 23-SEP-2000; 2000US-234567P.
 PR 25-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 28-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237395P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 XX
 PA (AVAL-) AVALON PHARM.
 XX
 PI Young PE, Augustus M, Carter KC, Ebner R, Endreess G, Horrigan S;
 PI Sopet DR, Weaver Z;
 XX
 DR WPI, 2002-188264/24.
 XX
 XX Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set -
 XX
 PS Claim 1; SEQ ID 6922; 44pp; English.
 XX
 XX The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 847 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytoskeletal
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.
 XX
 SQ Sequence 4797 BP; 957 A; 1349 C; 1366 G; 1099 T; 6 other;
 Query Match 87.3%; Score 285.8; DB 24; Length 4797;
 Best Local Similarity 96.1%; Pred. No. 1.6e-62;
 Matches 293; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 4 GAGGGGCGCAGGGGCTTGAGATCTCTCTGACAGACCGCCGCTCTGCTGAGGCGG 63
 Db 3765 GAGGGGCGCAGGGGCTTGAGATCTCTCTGACAGACCGCCGCTCTGCTGAGGCGG 3824
 QY 64 TCTCCAGGGGCTGCTCTCTCTGAGAAATGACAGAGGGGTGTCTTGGGACAGCTGCTCT 123
 Db 3825 TCTCCAGGGGCTGCTCTCTCTGAGAAATGACAGAGGGGTGTCTTGGGACAGCTGCTCT 3884
 QY 124 GAGGGGCTCTCCATCCAGGCGCAGGTTCTCGCTTACCTCTGAGCCCGACCCCTGGGCGCTG 183
 Db 3885 GAGGGGCTCTCCATCCAGGCGCAGGTTCTCGCTTACCTCTGAGCCCGACCCCTGGGCGCTG 3944
 QY 184 GGGTGGAAATCAGGAATATTTTCCAAAGAGTGAATGCTTTTGGGAAACCTCTAC 243
 Db 3945 GGGTGGAAATCAGGAATATTTTCCAAAGAGTGAATGCTTTTGGGAAACCTCTAC 4004
 QY 244 TTAATCCAAATGGGTTTCTCTCTGACAGTGAATTTCCAAATGTAATTAATTAATA 303
 Db 4005 TTAATCCAAATGGGTTTCTCTCTGACAGTGAATTTCCAAATGTAATTAATTAATA 4064
 QY 304 AAGTA 308
 Db 4065 AAGTA 4069

RESULT 8

ABL49089
 ID ABL49089 standard, cDNA; 4797 BP.

AC ABL49089;

DT 01-MAY-2002 (first entry)

DE Human lung tumour cDNA sequence for contig 31 SEQ ID NO:134.

KW Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;

KW immune response; ss.

OS Homo sapiens.

PN WO200200174-A2.

PD 03-JAN-2002.

PF 28-JUN-2001; 2001WO-US21065.

PR 28-JUN-2000; 2000US-0606421.

PR 02-AUG-2000; 2000US-0630940.

PR 21-AUG-2000; 2000US-0643597.

PR 15-SEP-2000; 2000US-0662766.

PR 09-OCT-2000; 2000US-0685696.

PR 12-DEC-2000; 2000US-0735705.

PR 07-MAY-2001; 2001US-0850716.

PA (CORI-) CORIXA CORP.

PI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;

PI McNeill PD, Fanger N, Retter MW, Marnierakis M, Fanger GR;

PI Vedvick TS, Carter D, Matanabe Y, Pecham DW;

DR WPI; 2002-090513/12.

PS Example 3; Page 238-240; 374pp; English.

CC The present invention describes human lung tumour proteins. Human lung

CC tumour proteins and polynucleotides have cytostatic and immunostimulant

CC activities, and can be used in vaccine production. Compositions

CC comprising the lung tumour proteins, polynucleotides, antibodies,

CC fusion proteins, T cell populations, or antigen presenting cells that

CC express the lung tumour proteins are useful for treating lung cancer or

CC stimulating an immune response. ABL49859 to ABL49300 and ABL74946 to

CC ABL75070 represent sequences used in the exemplification of the present

CC invention.

XX

SO Sequence 4797 BP; 957 A; 1349 C; 1386 G; 1099 T; 6 other;

Query Match 87.9%; Score 285.8; DB 24; Length 4797;

Best Local Similarity 96.1%; Pred. No. 1,6e-62;

Matches 293; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 4 GAGGGGCGCAGGGGCTTGAGATCTCTCTGACAGACCGCCGCTCTGCTGAGGCGG 63

Db 3765 GAGGGGCGCAGGGGCTTGAGATCTCTCTGACAGACCGCCGCTCTGCTGAGGCGG 3824

QY 64 TCTCCAGGGGCTGCTCTCTCTGAGAAATGACAGAGGGGTGTCTTGGGACAGCTGCTCT 123

Db 3825 TCTCCAGGGGCTGCTCTCTCTGAGAAATGACAGAGGGGTGTCTTGGGACAGCTGCTCT 3884

QY 124 GAGGGGCTCTCCATCCAGGCGCAGGTTCTCGCTTACCTCTGAGCCCGACCCCTGGGCGCTG 183

Db 3885 GAGGGGCTCTCCATCCAGGCGCAGGTTCTCGCTTACCTCTGAGCCCGACCCCTGGGCGCTG 3944

QY 184 GGGTGGAAATCAGGAATATTTTCCAAAGAGTGAATGCTTTTGGGAAACCTCTAC 243

Db 3945 GGGTGGAAATCAGGAATATTTTCCAAAGAGTGAATGCTTTTGGGAAACCTCTAC 4004

QY 244 TTAATCCAAATGGGTTTCTCTCTGACAGTGAATTTCCAAATGTAATTAATTAATA 303

Db 4005 TTAATCCAAATGGGTTTCTCTCTGACAGTGAATTTCCAAATGTAATTAATTAATA 4064

QY 304 AAGTA 308

Db 4065 AAGTA 4069

RESULT 9

AAT23809

ID AAT23809 standard; cDNA to mRNA; 287 BP.

AC AAT23809;

DT 15-AUG-1996 (first entry)

DE Human gene signature H0WG505718.

KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;

KW human; cloning; mapping; non-biased library; diagnosis; detection;

OS Homo sapiens.

PN WO9514772-A1.

PD 01-JUN-1995.

PF 11-NOV-1994; 94WO-JP01916.

PR 12-NOV-1993; 93JP-0355504.

PA (MATS/) MATSUBARA K.

PI (OKUB/) OKUBO K.

DR WPI; 1995-206931/27.

PS Identifying gene signatures in 3'-directed human cDNA library - e.g.

PS for diagnosis of abnormal cell function, by preparing cDNA that

PS reflects relative abundance of corresp. mRNA in specific human

PS tissues

PS Claim 1; Page 1460; 2245pp; Japanese.

XX A single-stranded DNA (or its complementary strand or the corresp.
 CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
 CC given in AAT19901-T26837 and which is able to hybridise to part of
 CC human genomic DNA, cDNA or mRNA is claimed. The GS (gene Signature)
 CC sequences were obtained from 3'-directed cDNA libraries prepared
 CC from various human tissues; synthesis of cDNA was initiated from the
 CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
 CC untranslated sequence is unique to a particular mRNA species, almost
 CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
 CC is constructed so as to reflect accurately the relative abundance of
 CC different mRNAs in the particular tissue from which it was derived.
 CC The appearance frequency of a given GS in a cDNA library can be
 CC determined (esp. using primers and probes derived from the GS
 CC sequences) as a means of diagnosing abnormal cell function or for
 CC recognising different cell types.
 XX

Sequence 287 BP; 60 A; 78 C; 67 G; 80 T; 2 other;

Query Match 86.2%; Score 280.2; DB 16; Length 287;
 Best Local Similarity 98.3%; Pred. No. 1.7e-61;
 Matches 282; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 24 GATCCTCTCCAGACCCAGCCCGCTCTGCTGAGCGCTCCAGGAGGCTGCTCTC 83
 Db 1 GATCCTCTCCAGACCCAGCCCGCTCTGCTGAGCGCTCCAGGAGGCTGCTCTC 60
 QY 84 CTGGAATTTAGACAGGGGCTGCTTGGGAGAGCTGAGCTCTGAGCGCTCCATCCAGGCC 143
 Db 61 CTGGAATTTAGACAGGGGCTGCTTGGGAGAGCTGAGCTCTGAGCGCTCCATCCAGGCC 120
 QY 144 AGCTTCCGCTAGCTCCCTGCTGAGCGCTCCAGCGCTCCAGGAGGCTGCTCTC 203
 Db 121 AGCTTCCGCTAGCTCCCTGCTGAGCGCTCCAGCGCTCCAGGAGGCTGCTCTC 180
 QY 204 TCCAAAGAGTAGTCTTTTCTTGGCAAACTCTACTTAATCCAAATGGTTTCT 263
 Db 121 TCCAAAGAGTAGTCTTTTCTTGGCAAACTCTACTTAATCCAAATGGTTTCT 240
 QY 264 CTGTACAGTAGATTTTCCAAATGTATTAACCTTTAATATTAAGTAAA 310
 Db 241 CTGTACAGTAGATTTTCCAAATGTATTAACCTTTAATATTAAGTAAA 287

RESULT 10
 AAT24631
 ID AAT24631 standard; cDNA; 4787 BP.
 XX
 AC AAT24631;
 XX
 DT 07-DEC-1999 (first entry)
 XX
 DE Human lung tumor associated polynucleotide.
 XX
 KW Human; lung tumor; lung cancer; T cell stimulation; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO9947674-A2.
 XX
 PD 23-SEP-1999.
 XX
 PF 17-MAR-1999; 99WO-US05798.
 XX
 PR 18-MAR-1998; 98US-0040802.
 PR 18-MAR-1998; 98US-0040984.
 PR 27-JUL-1998; 98US-0123912.
 PR 27-JUL-1998; 98US-0123933.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Read SG, Wang T;
 XX

DR WPI; 1999-571839/48.

XX
 PT New isolated lung tumor polynucleotides, used to develop products for
 PT the treatment, prevention and monitoring the progression of lung cancer
 PT -
 PS Claim 12; Page 122-123; 148pp; English.
 XX

CC The invention provides isolated human lung tumor nucleic acids and
 CC polypeptides. The polypeptides can be used for the treatment of lung
 CC cancer. The polypeptides and polynucleotides can be used to stimulate T
 CC cells or antigen presenting cells for use in the treatment of lung
 CC cancer. The polypeptides and monoclonal antibodies specific for the
 CC polypeptides can also be used to inhibit the development of lung cancer.
 CC Agents which bind the polypeptides can be used for detecting lung cancer
 CC and for monitoring the progression of lung cancer.
 XX

Sequence 4787 BP; 957 A; 1346 C; 1386 G; 1092 T; 6 other;

Query Match 83.0%; Score 269.8; DB 20; Length 4787;
 Best Local Similarity 95.1%; Pred. No. 1.8e-58;
 Matches 290; Conservative 0; Mismatches 12; Indels 3; Gaps 1;

QY 4 GAGGCGGCGAGGGGCTGAGAGATCCTCTGAGACACAGCCGCTGCTGAGCGCG 63
 Db 3764 GAGGCGGCGAGGGGCTGAGAGATCCTCTGAGAGCTCAGCCCGCTCTGAGCGCG 3823
 QY 64 TCTCCAGGGGCTCTCTCTCTGGAATTTGACAGGGGCTCTTGGGAGAGCTGCTCT 123
 Db 3824 TCTCCAGGGGCTCTCTCTCTGGAATTTGACAGGGGCTCTTGGGAGAGCTGCTCT 3883
 QY 124 GAGGCGCTCCATCCAGGCGAGGCTCTCCCTGAGCTCTCTGAGCGCGCTG 183
 Db 3884 GAGGCGCTCCATCCAGGCGAGGCTCTCCCTGAGCTCTCTGAGCGCGCTG 3943
 QY 184 GAGTGAATGAGATATTTTCCAAAGAGTAGTACTTTTCTTTGGCAAACTCTAC 243
 Db 3944 GAGTGAATGAGATATTTTCCAAAGAGTAGTACTTTTCTTTGGCAAACTCTAC 4003
 QY 244 TTAATCCAAATGGTTTTTCTCTGTCAGTAGATTTTCCAAATGTATTAACCTTAAATA 303
 Db 4004 TTAATCCAAATGG--TTCTCTGTACAGTAGATTTTCCAAATGTATTAACCTTAAATA 4060
 QY 304 AAGTA 308
 Db 4061 AAGTA 4065

RESULT 11
 AAT72270
 ID AAT72270 standard; DNA; 425 BP.
 XX
 AC AAT72270;
 XX
 DT 15-APR-2002 (first entry)
 XX
 DE Exons 2-6 of Syndecan 1 gene.
 XX
 KW Lung; cancer; metastasis; solid tumour; blood; bone marrow; syndecan 1;
 KW collagen 1 alpha 2; 7013; 7018; amplification; mammal; human; dog; cat;
 KW bile duct; colon; breast; uterus; oesophagus; larynx; liver; brain;
 KW remission; relapse; gene; ss.
 XX
 OS Synthetic.
 XX
 PN WO200198539-A2.
 XX
 PD 27-DEC-2001.
 XX
 PF 21-JUN-2001; 2001WO-US19980.
 XX
 PR 21-JUN-2000; 2000US-215727P.
 PR 27-OCT-2000; 2000US-243976P.
 XX

XX (HITB) HITACHI CHEM CO LTD.
 PA (HITB) HITACHI CHEM RES CENT INC.
 PA (HITB) HITACHI LTD.
 XX Mitsuhashi M, Kambara H, Matsumaga H, Kawamura M;
 XX MPI; 2002-098233/13.
 DR
 XX
 PT Identifying lung cancer/metastasis of solid tumor in patient by
 PT isolating blood or non-lung tissue, or bone marrow from patient and
 PT identifying presence of marker e.g. syndecan 1, collagen 1 alpha 2,
 PT 7013, or 7018 -
 XX
 XX Example 2, Fig 6, 29pp; English.
 PS
 XX This sequence corresponds to exons 2-6 of the Syndecan 1 gene.
 CC Syndecan 1 is a cell surface transmembrane heparan sulphate proteoglycan
 CC from the family of proteoglycans that bind to extracellular matrix and
 CC growth factors. The loss of regulation of this gene has been identified
 CC in several cancers. This sequence was identified using the method of
 CC the invention for identifying lung cancer or metastasis of a solid
 CC tumour. The method comprises isolating blood (or non-lung tissue in
 CC the case of identifying lung cancer, or bone marrow in case of
 CC identifying metastasis) from a patient, and identifying the presence
 CC of at least one marker (M) such as syndecan 1, collagen 1 alpha 2, 7013,
 CC or 7018. This sequence was found to be more abundant in lung cancer
 CC RNA than in normal blood. The method of the invention is useful for
 CC identifying lung cancer in a mammal e.g., human, dog or cat, and
 CC identifying metastasis of solid tumour in a patient, where the solid
 CC tumour is of bile duct, colon, breast, uterus, oesophagus or larynx. The
 CC method is useful for identifying presence of lung cancer cells in the
 CC blood or bone marrow, and also for identifying metastasis and thus for
 CC identifying lung cancer cells in an organ such as liver or brain. The
 CC method is useful to identify the presence of lung cancer cells at a
 CC very early stage in the disease, or after remission or to identify a
 CC relapse.
 CC
 XX Sequence 425 BP; 62 A; 124 C; 146 G; 93 T; 0 other;
 SQ
 Query Match 55.8%; Score 181.4; DB 24; Length 425;
 Best Local Similarity 94.5%; Pred. No. 1.9e-36;
 Matches 188; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
 QY 4 GAGGCGGCGAGGCGCTGGAGATCTCTCGAGACAGCCCGCTGCTGCGCGCG 63
 Db 227 GAGGCGGCGAGGCGCTGGAGATCTCTCGAGGCTCAAGCCCGCTGCTGCGCGCG 286
 QY 64 TCTCCAGGCGGCTGCTCTCTCGAAATGACGAGGCGTCTTGGGCGAGAGTGGCTCT 123
 Db 287 TCTCCAGGCGGCTGCTCTCTCGAAATGACGAGGCGTCTTGGGCGAGAGTGGCTCT 346
 QY 124 GAGCGCTTCATCCAGAGGCTCTCCGTAGCTCTGTGCGCCCACTGCGGCGCTG 183
 Db 347 GAGCGCTTCATCCAGAGGCTCTCCGTAGCTCTGTGCGCCCACTGCGGCGCTG 406
 QY 184 GCGTGAATCAGGAATATT 202
 Db 407 GCGTGAATCAGGAATATT 425

RESULT 12
 AAK53003/c
 ID AAK53003 standard; cDNA; 492 BP.

XX AAK53003;
 XX
 XX 21-MAY-2002 (first entry)
 XX
 DE cDNA encoding lung tumour protein clone R0128.E06.
 XX
 XX Lung tumour; cancer; T cell; immune response stimulator;
 KM cytoskeletal; gene; ss.

XX OS Homo sapiens.
 XX
 PN WO200204514-A2.
 XX
 XX 17-JAN-2002.
 PD
 XX
 PF 10-JUL-2001; 2001WO-US22058.
 XX
 XX 11-JUL-2000; 2000US-0614124.
 PR 29-AUG-2000; 2000US-0651563.
 PR 08-SEP-2000; 2000US-0658824.
 PR 26-SEP-2000; 2000US-0671325.
 PR 06-OCT-2000; 2000US-0677419.
 PR 30-OCT-2000; 2000US-0702705.
 PR 13-DEC-2000; 2000US-0736457.
 PR 03-MAY-2001; 2001US-0849626.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Wang T, Matanabe Y, Henderson RA, Johnson JC, Retter MW,
 PI Marnerakis M, Carter D, Fanger GR, Vedvick TS, Bangur CS,
 PI Mcnabb A, Wang A, Fanger N, Switzer A, Monelli PD, Clapper JD;
 XX MPI; 2002-164634/21.
 DR
 XX
 PT Novel polynucleotide encoding a lung tumour polypeptide useful for
 PT stimulating and/or expanding T cells specific for a tumour protein -
 PS
 XX Claim 1; SEQ ID No 1041; 223pp; English.
 CC The invention describes an isolated polynucleotide and polypeptide
 CC useful for stimulating and/or expanding T cells specific for a tumour
 CC protein for determining the presence of a cancer in a patient. A
 CC composition containing the polynucleotide and/or polypeptide is useful
 CC for treating a lung cancer in a patient. The polypeptide is useful for
 CC removing tumour cells from a biological sample. The polynucleotide is
 CC also useful as probe or primer to detect the level of mRNA encoding a
 CC tumour protein. This sequence encodes a lung tumour associated protein
 CC or protein fragment, described in the method of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 492 BP; 153 A; 126 C; 110 G; 102 T; 1 other;
 SQ
 Query Match 44.4%; Score 144.4; DB 24; Length 492;
 Best Local Similarity 98.6%; Pred. No. 4.7e-27;
 Matches 145; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 162 TGTGCCCCACCCCTGGCGCTGGAATCAGAAATATTTCCAAAGATGATGCT 221
 Db 492 TGTGCCCCACCCCTGGCGCTGGAATCAGAAATATTTCCAAAGATGATGCT 433
 QY 222 TTTGCTTTTGGCAAAATCTACTTAATCCATGSGTTTCTCTGTACAGTATTTCC 281
 Db 432 TTTGCTTTTGGCAAAATCTACTTAATCCATGSGTTTCTCTGTACAGTATTTCC 373
 QY 282 AAATGTAATAAATTTAATAATAAGTA 308
 Db 372 AAATGTAATAAATTTAATAATAAGTA 346

RESULT 13
 AAK54184/c
 ID AAK54184 standard; cDNA; 286 BP.

XX AAK54184;
 XX
 XX 16-NOV-2001 (first entry)
 XX
 DE Murine transport and binding associated protein encoding cDNA SEQ ID 749.

XX	Murine; liver; gene library; amino acid synthesis; binding protein;
KX	cell metabolism; energy metabolism; fatty acid metabolism; synthesis;
KW	phospholipid metabolism; purine; pyrimidine; nucleoside; nucleotide;
KV	replication; transcription; translation; transport protein; ss.
XX	
OS	Mus musculus.
XX	
PD	DE20103510-U1.
PN	
XX	
PD	07-JUN-2001.
XX	
PF	28-FEB-2001; 2001DE-2003510.
XX	
FR	02-DEC-1999, 99DE-1058160.
XX	
PA	(LION-) LION BIOSCIENCE AG.
XX	
DR	WPI; 2001-368570/39.
XX	
PT	Gene library containing sequences with specific 3'-ends and no polyA
PT	tail, encoding proteins involved in a wide range of cellular processes
PT	-
PS	Claim 15; Page 227; 25ipp; German.
XX	
CC	This invention describes a novel gene library (A) comprises a gene
CC	sequence (or its part) encoding a protein involved in amino acid
CC	synthesis, cellular/energy metabolism, metabolism of
CC	fatty acids/phospholipids, synthesis or breakdown of
CC	purines/pyrimidines/nucleosides/nucleotides, DNA
CC	replication/transcription/translation, or is a transport/binding protein.
CC	(A) are produced that correspond to the 3'-end of mRNA but without the
CC	polya tail. They can be prepared more efficiently and with less effort
CC	than conventional libraries. AK53436-AK54275 represent fragments of the
CC	gene library described in the method of the invention.
XX	
SQ	Sequence 286 BP; 78 A; 57 C; 74 G; 77 T; 0 other;
	Query Match 38.0%; Score 123.4; DB 22; Length 286;
	Best Local Similarity 76.2%; Pred No. 8e-22;
	Matches 208; Conservative 0; Mismatches 51; Indels 14; Gaps 4
OY	36 GAACAGCCCCGTCCTGTGGCGCCGTCACAGGSGCTGCTCCTCGNAATTGAC 95
DB	279 GCCCAAGCGTCCACCACCTTTGGAACATCTCTAGTCATCTTCTCCCGGAAGTTAC 220
OY	96 GAGGGGTGCTTGGGCAGAGTGCGTCTGAGCGCTCCATCCAAAGCCAGGTTCTCCGTT 155
DB	219 AAACAACATCTGAGTAGTGCTGCGACTG-GTTCCTCCATCAAGAACCAGTTCACCTTC 161
OY	156 AGCTCTGTGGGCCCCCACCCCTGGGCGCGGTGGAATCAGGAATATTTTCCAAAGAGTGA 215
DB	160 AGCTCTGTGGGCCCC-----GCCCCAGGCTGGAGTGAAGATGTTTCCCAAAGAGTG- 108
OY	216 TAGCTTTTGTGCTTTGGCAAACTACTTAATCCAATGAGGTTTTTCTCTGACAGTAGA 275
DB	107 -AGCTTTTGTGCTTTGGCAAAAGCGACTTAATCCAAGGCT-----TCTGACAGTGA 54
OY	276 TTTTCCAAATGTATAAACTTTAATATAAGTA 308
DB	53 TTTTGCAGATGTATAAACTTTAATATAAGTA 21
	RESULT 14
ID	AAQ67902
ID	AAQ67902 standard; DNA; 26700 BP.
AC	AAQ67902;
DT	08-DEC-1994 (first entry)
DB	Syndecan gene.

[illegible]

Fri Jan 10 09:34:16 2003

us-09-825-682a-57.rng

Page 11

Search completed: January 10, 2003, 01:54:11
Job time : 127.067 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 10, 2003, 01:47:01 ; Search time 18.99 Seconds
(without alignments)
5248.554 Million cell updates/sec

Title: US-09-825-682A-57

Perfect score: 325
Sequence: 1 aaagagggcgcgcagggcct.....gtaaaaaaaaaaaaaaaa 325

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 segs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES			
Result	Score	Query Match Length	ID
No.			
1	307.4	94.6	2430 2 US-08-488-199-3
2	285.8	87.9	4797 4 US-09-643-597-134
3	113	34.8	26700 1 US-08-472-217-1
4	113	34.8	26700 2 US-08-488-199-5
5	113	34.8	26700 3 US-08-760-534A-1
6	112.6	34.6	2432 1 US-08-078-683A-1
7	38.8	11.9	958 2 US-08-757-046A-5
8	38.8	11.9	958 3 US-09-447-208-5
9	38.8	11.9	958 4 US-09-135-988-5
10	38.8	11.9	958 4 US-09-277-715-5
11	38.8	11.9	958 4 US-08-597-274A-5
12	38.8	11.9	958 4 US-08-908-909-5
13	38.8	11.9	958 4 US-09-609-161B-5
14	38.8	11.9	958 4 US-08-990-103-5
15	38.4	11.8	8920 2 US-08-446-855A-1
16	38.4	11.8	8920 4 US-09-150-741-1
17	38.2	11.8	1332 2 US-09-057-762-1
18	38.2	11.8	1332 3 US-08-326-119A-1
19	37.6	11.6	1260 1 US-08-599-252-79
20	37.6	11.6	1260 1 US-08-436-074-52
21	37.6	11.6	1260 1 PCT-US96-06352-79
22	37.6	11.6	1260 5 PCT-US96-06583-79
23	37	11.4	988 1 US-08-684-862-10
24	36.4	11.2	579 4 US-09-040-984-75
25	36.4	11.2	579 4 US-09-123-912-75
26	36.4	11.2	579 4 US-09-643-597-75
27	36.2	11.1	4257 2 US-08-690-473-1

28	36.2	11.1	4257 4 US-09-259-821A-1	Sequence 1, Appl1
29	36.2	11.1	4257 4 US-08-843-659-1	Sequence 1, Appl1
30	36.2	11.1	4337 3 US-09-187-049-1	Sequence 1, Appl1
c	31	36.2	12001 1 US-08-458-568B-11	Sequence 11, Appl1
32	35.4	10.9	2851 4 US-09-535-521-1	Sequence 1, Appl1
33	35.4	10.9	2851 4 US-09-535-521-3	Sequence 3, Appl1
34	35.4	10.9	3785 1 US-08-445-640-9	Sequence 9, Appl1
35	35.4	10.9	3785 3 US-08-170-558-9	Sequence 9, Appl1
36	35.4	10.9	3785 3 US-08-447-314-9	Sequence 9, Appl1
37	35.2	10.8	3785 3 US-08-445-461-9	Sequence 9, Appl1
38	35.2	10.8	1172 1 US-07-945-288-9	Sequence 9, Appl1
39	35.2	10.8	1172 1 US-08-462-831-9	Sequence 9, Appl1
40	35.2	10.8	1172 1 US-08-461-809-9	Sequence 9, Appl1
41	35.2	10.8	1172 1 US-08-461-441-9	Sequence 9, Appl1
42	35.2	10.8	1172 5 PCT-US93-08518-9	Sequence 9, Appl1
43	35.2	10.8	1841 5 PCT-US95-00362-1	Sequence 1, Appl1
44	35	10.8	3214 1 US-08-484-105-17	Sequence 17, Appl1
45	35	10.8	3214 1 US-08-484-106-17	Sequence 17, Appl1

ALIGNMENTS

```
RESULT 1
US-08-488-199-3
; Sequence 3, Application US/08488199
; Patent No. 5851993
;
; GENERAL INFORMATION:
; APPLICANT: Valkenen, Markku
; TITLE OF INVENTION: Suppression of Tumor Cell Growth By
; TITLE OF INVENTION: Syndecan-1 Ectodomain
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX
; STREET: 1100 New York Ave., NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,199
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/258,862
; FILING DATE: 13-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbal, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 1102.0130001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2430 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 206..1138
; US-08-488-199-3
;
; Query Match 94.6%; Score 307.4; DB 2; Length 2430;
; Best Local Similarity 99.1%; Pred. No. 1.7e-73;
; Matches 319; Conservative 0; Mismatches 2; Indels 1;
```

```
QY      4  GAGGCGGAGGGGCTGAGATCTCTCGACAGACACGCCGTCGCTGTGGGCGCG 63
          |||
Db      2100  GAGGCGGAGGGGCTGAGATCTCTCGACAGACACGCCGTCGCTGTG-6GCGG 2158
QY      64  TCTCCAGGGGCTGCTTCTCTCGGAAATTGACAGGGGCTGCTTGAGGACAGCTGCTT 123
          |||
Db      2159  TCTCCAGGGGCTGCTTCTCTCGGAAATTGACAGGGGCTGCTTGAGGACAGCTGCTT 2218
QY      124  GAGCGCTTCATCCAGGCGCAGGTTCTCCGTTAGCTCTGTGGGCCCACTTGGGCGCTG 183
          |||
Db      2219  GAGCGCTTCATCCAGGCGCAGGTTCTCCGTTAGCTCTGTGGGCCCACTTGGGCGCTG 2278
QY      184  GGCTGGAATCGAGAAATATTTTCCAAAGAGTGAATGTTTTCCTTTGGGAAAACCTTAC 243
          |||
Db      2279  GGCTGGAATCGAGAAATATTTTCCAAAGAGTGAATGTTTTCCTTTGGGAAAACCTTAC 2338
QY      244  TTAATCAATGGGTTTTCTCTGTACAGTAGATTTTCCAAATGTAATTAATTATATA 303
          |||
Db      2339  TTAATCAATGGGTTTTCTCTGTACAGTAGATTTTCCAAATGTAATTAATTATATA 2398
QY      304  AAGTAAAAAAAAAAAAAAAA 325
          |||
Db      2399  AAGTAAAAAAAAAAAAAAAA 2420

RESULT 2
US-09-643-597-134
; Sequence 134, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Pan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fager, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643.597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 134
; LENGTH: 4797
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(4797)
; OTHER INFORMATION: n = A,T,C or G
US-09-643-597-134

Query Match      87.9%; Score 285.8; DB 4; Length 4797;
Best Local Similarity 96.1%; Pred. No. 1.4e-67;
Matches 293; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
```

```
Db      3885  GAGCGCTTCATCCAGGCGCAGGTTCTCCGTTAGCTCTGTGGGCCCACTTGGGCGCTG 3944
QY      184  GGCTGGAATCGAGAAATATTTTCCAAAGAGTGAATGTTTTCCTTTGGGAAAACCTTAC 243
          |||
Db      3945  GGCTGGAATCGAGAAATATTTTCCAAAGAGTGAATGTTTTCCTTTGGGAAAACCTTAC 4004
QY      244  TTAATCAATGGGTTTTCTCTGTACAGTAGATTTTCCAAATGTAATTAATTATATA 303
          |||
Db      4005  TTAATCAATGGGTTTTCTCTGTACAGTAGATTTTCCAAATGTAATTAATTATATA 4064
QY      304  AAGTA 308
          |||
Db      4065  AAGTA 4069

RESULT 3
US-08-472-217-1
; Sequence 1, Application US/08472217
; Patent No. 5726058
; GENERAL INFORMATION:
; APPLICANT: Alanen-Kurki, Leena
; APPLICANT: Auvinen, Petri
; APPLICANT: Uakkola, Panu
; APPLICANT: Jalonen, Markku
; APPLICANT: Lepp, Sirpa
; APPLICANT: Maki, Markku
; APPLICANT: Viuhinen, Tapani
; APPLICANT: W rti, Anni
; TITLE OF INVENTION: Syndecan Stimulation Of Cellular
; TITLE OF INVENTION: Differentiation
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472.217
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/206,186
; FILING DATE: 07-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/988,427
; FILING DATE: 01-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbala, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 1102.0050003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26700 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(4378..4443, 22026..22106, 23001..23483,
; LOCATION: 23905..24039, 24251..24418)
```



```
/
/   REGISTRATION NUMBER: 33,851
/   REFERENCE/DOCKET NUMBER: 1708.0050004/MAC
/   TELECOMMUNICATION INFORMATION:
/   TELEPHONE: (202) 371-2600
/   TELEFAX: (202) 371-2540
/   INFORMATION FOR SEQ ID NO: 1:
/   SEQUENCE CHARACTERISTICS:
/   LENGTH: 26700 base pairs
/   TYPE: nucleic acid
/   STRANDEDNESS: double
/   TOPOLOGY: linear
/   MOLECULE TYPE: DNA (genomic)
/   FEATURE:
/   NAME/KEY: CDS
/   LOCATION: join(4378..4443,22026..22106,23001..23483,
/   LOCATION: 23905..24039,24251..24418)
/   US-08-760-534A-1

Query Match      34.8%; Score 113; DB 3; Length 26700;
Best Local Similarity 76.2%; Pred. No. 6.7e-21;
Matches 208; Conservative 0; Mismatches 50; Indels 15; Gaps 5;

QY 36 GACCACGCCCGCTGCTGCTGCGCGCGCTCCAGGGGCTGCTCTCTGGAATTGAC 95
DB 25419 GCCCAAGCGTCCCGACCTTTGTACATCTCTA-GTCATCTTCTCCCGAAGTTGAC 25477

QY 96 GAGGGGCTGTGGGAGAGCTGGTCTGAGCGCTTCCATCCAGGCCAGTTTCCGTT 155
DB 25478 AAGACACATCTTGAGTATGCTGGGACTG-GTCTCCATCAAGAACCAATTCCACTTC 25536

QY 156 AGCTCCTGTGGCCCGCCACCTGCGGCGTGAATCAGAAATTTTCCAAAGATGA 215
DB 25537 AGCTCCTGTGGCCCC-----GCCCCAGGCTGAGTCAAAATTTTCCAAAGATG- 25589

QY 216 TAGCTTTTCTTTGGCAAACTGACTTAATCCATGAGTTTCTCTGTACAGTAGA 275
DB 25590 -AGCTTTTCTTTGGCAAAAGCTACTTAATCAATGGGT-----TCTGTACAGTAGA 25643

QY 276 TTTTCCAAATGTAACTTAATATTAAGTA 308
DB 25644 TTTTGACAGATGTAACTTAATATTAAGTA 25676

RESULT 6
US-08-078-683A-1
/   Sequence 1, Application US/08078683A
/   Patent No. 5486599
/   GENERAL INFORMATION:
/   APPLICANT: Saunders, Scott
/   APPLICANT: Bernfield, Meiron
/   APPLICANT: Kato, Masato
/   TITLE OF INVENTION: Construction and Use of Synthetic
/   NUMBER OF SEQUENCES: 43
/   CORRESPONDENCE ADDRESS:
/   ADDRESSEE: LAHIVE & COCKFIELD
/   STREET: 60 State Street
/   CITY: Boston
/   STATE: MA
/   COUNTRY: USA
/   ZIP: 02109
/   COMPUTER READABLE FORM:
/   MEDIUM TYPE: Floppy disk
/   OPERATING SYSTEM: PC-DOS/MS-DOS
/   SOFTWARE: ASCII (text)
/   CURRENT APPLICATION DATA:
/   APPLICATION NUMBER: US/08/078.683A
/   FILING DATE: 17-JUN-1993
/   CLASSIFICATION: 435
/   ATTORNEY/AGENT INFORMATION:
/   NAME: Vincent, Matthew P.
/   REGISTRATION NUMBER: 36,709
```

```
/
/   REFERENCE/DOCKET NUMBER: CME-062
/   TELECOMMUNICATION INFORMATION:
/   TELEPHONE: (617) 227-7400
/   TELEFAX: (617) 227-5941
/   INFORMATION FOR SEQ ID NO: 1:
/   SEQUENCE CHARACTERISTICS:
/   LENGTH: 2432 base pairs
/   TYPE: nucleic acid
/   STRANDEDNESS: single
/   TOPOLOGY: linear
/   MOLECULE TYPE: cDNA
/   FEATURE:
/   NAME/KEY: CDS
/   LOCATION: 240..1175
/   FEATURE:
/   NAME/KEY: misc.feature
/   LOCATION: 305..306
/   OTHER INFORMATION: /function= "Exon 1/Exon2 boundary"
/   FEATURE:
/   NAME/KEY: misc.feature
/   LOCATION: 389..390
/   OTHER INFORMATION: /function= "Exon 2/Exon 3 boundary"
/   FEATURE:
/   NAME/KEY: misc.feature
/   LOCATION: 869..870
/   OTHER INFORMATION: /function= "Exon 3/Exon 4 boundary"
/   US-08-078-683A-1

Query Match      34.6%; Score 112.6; DB 1; Length 2432;
Best Local Similarity 76.4%; Pred. No. 3.5e-21;
Matches 207; Conservative 0; Mismatches 49; Indels 15; Gaps 5;

QY 36 GACCACGCCCGCTGCTGCTGCGCGCGCTCCAGGGGCTGCTTCTCTGGAATTGAC 95
DB 2176 GCCCAAGCGTCCCGACCTTTGTACATCTCTA-GTCATCTTCTCCCGAAGTTGAC 2234

QY 96 GAGGGGCTGTGGGAGAGCTGGTCTGAGCGCTTCCATCCAGGCCAGTTTCCGTT 155
DB 2235 AAGACACATCTTGAGTATGCTGGGACTG-GTCTCCATCAAGAACCAATTCCACTTC 2293

QY 156 AGCTCCTGTGGCCCGCCACCTGCGGCGTGAATCAGAAATTTTCCAAAGATGA 215
DB 2234 AGCTCCTGTGGCCCC-----GCCCCAGGCTGAGTCAAAATTTTCCAAAGATG- 2346

QY 216 TAGCTTTTCTTTGGCAAACTGACTTAATCCATGAGTTTCTCTGTACAGTAGA 275
DB 2347 -AGCTTTTCTTTGGCAAAAGCTACTTAATCAATGGGT-----TCTGTACAGTAGA 2400

QY 276 TTTTCCAAATGTAACTTAATATTAAG 306
DB 2401 TTTTGACAGATGTAACTTAATATTAAG 2431

RESULT 7
US-08-757-046A-5
/   Sequence 5, Application US/08757046A
/   Patent No. 5876995
/   GENERAL INFORMATION:
/   APPLICANT: Bryan, Bruce
/   APPLICANT: Brown, Martin
/   TITLE OF INVENTION: BIOLUMINESCENT ARTICLES OF MANUFACTURE
/   NUMBER OF SEQUENCES: 14
/   CORRESPONDENCE ADDRESS:
/   ADDRESSEE: Brown, Martin, Haller & McClain
/   STREET: 1660 Union Street
/   CITY: San Diego
/   STATE: CA
/   COUNTRY: USA
/   ZIP: 92101-2926
/   COMPUTER READABLE FORM:
/   MEDIUM TYPE: Diskette
/   OPERATING SYSTEM: DOS
/   SOFTWARE: FastSeq Version 1.5
```



```

EARLIER APPLICATION NUMBER: 60/102,939
EARLIER FILING DATE: 1998-10-01
EARLIER APPLICATION NUMBER: 60/089,367
EARLIER FILING DATE: 1998-06-15
EARLIER APPLICATION NUMBER: 60/079,624
EARLIER FILING DATE: 1998-03-27
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 958
TYPE: DNA
ORGANISM: Aequorea (luminescent jellyfish)
FEATURE:
NAME/KEY: CDS
LOCATION: (115)..(702)
FEATURE:
OTHER INFORMATION: Apoaequorin-encoding gene
PUBLICATION INFORMATION:
PATENT DOCUMENT NUMBER: 5,093,240
PATENT FILING DATE: 1987-10-08
PUBLICATION DATE: 1992-03-03
PUBLICATION INFORMATION:
AUTHORS: Imouye, S.
TITLE: Cloning and sequence analysis of cDNA for the luminescent protein
JOURNAL: Proc. Natl. Acad. Sci. USA
VOLUME: 82(110)
PAGES: 3154-3158
DATE: 1985-05
US-09-277-716-5

Query Match      11.9%; Score 38.8; DB 4; Length 958;
Best Local Similarity 58.8%; Pred. No. 0.16;
Matches 67; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 212 GGAAGAGCTTTGGCTTTGGCAACCTACTTAATCCATGAGGGTTTCTGTGACAG 271
Db 812 GGTGTGATTTTGTGAATTTGGAACAGATTAAATGAGATTGATTTGTTTAAATCAA 871
QY 272 TAGATTTTCCAAATGTATATTAATTAATTAAGTAAAAAAGTAAAAAAGTAAAAA 325
Db 872 CAGAGCTTACAAATCGAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAA 925

RESULT 11
US-08-597-274A-5
Sequence 5, Application US/08597274A
Patent No. 6247995
GENERAL INFORMATION:
APPLICANT: Bryan, Bruce
TITLE OF INVENTION: BIOLUMINESCENT NOVELTY ITEMS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: CA
COUNTRY: USA
ZIP: 92101-2926
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/597,274A
FILING DATE: 02/06/96
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779

```

REFERENCE/DOCKET NUMBER: 6680-105
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-238-0999
TELEFAX: 619-238-0062
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 115...702
OTHER INFORMATION: apoaeguorin-encoding gene
PUBLICATION INFORMATION:
DOCUMENT NUMBER: 5,093,240
AUTHORS: Inouye et al.
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
VOLUME: 82
PAGES: 3154-3158
DATE: (1985)
US-08-597-274A-5

Query Match 11.9%; Score 38.8; DB 4; Length 958;
Best Local Similarity 58.8%; Pred. No. 0.16;
Matches 67; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 212 GGTATGCTTTTGGTCTTTGGCAAACTCTACTTAATCCAAATGGGTTTCTCTGACG 271
DB 812 GGTGATTTTGTATTTAGGAACAGATTAAATCGAATGATTCTGTTTTTATATCA 871

QY 272 TAGATTTCCAAATGTAATACTTTATATTAAGTAAAAA 325
DB 872 CAGAACTTCAAAATCGAAAAAGTAAAAA 925

RESULT 12
US-08-908-909-5
Sequence 5, Application US/08908909
Patent No. 6416960
GENERAL INFORMATION:
APPLICANT: Bryan, Bruce
TITLE OF INVENTION: DETECTION AND VISUALIZATION OF
TITLE OF INVENTION: NEOPLASTIC TISSUES AND OTHER TISSUES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: CA
COUNTRY: USA
ZIP: 92101-2926
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FaetsSO Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/908,909
FILING DATE: 08-AUG-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/023,374
FILING DATE: 08-AUG-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L

REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6680-108
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-238-0999
TELEFAX: 619-238-0062
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 115...702
OTHER INFORMATION: apoaeguorin-encoding gene
PUBLICATION INFORMATION:
AUTHORS: Inouye et al.
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
VOLUME: 82
PAGES: 3154-3158
DATE: (1985)
US-08-908-909-5

Query Match 11.9%; Score 38.8; DB 4; Length 958;
Best Local Similarity 58.8%; Pred. No. 0.16;
Matches 67; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 212 GGTATGCTTTTGGTCTTTGGCAAACTCTACTTAATCCAAATGGGTTTCTCTGACG 271
DB 812 GGTGATTTTGTATTTAGGAACAGATTAAATCGAATGATTCTGTTTTTATATCA 871

QY 272 TAGATTTCCAAATGTAATACTTTATATTAAGTAAAAA 325
DB 872 CAGAACTTCAAAATCGAAAAAGTAAAAA 925

RESULT 13
US-09-609-161B-5
Sequence 5, Application US/09609161B
Patent No. 643682
GENERAL INFORMATION:
APPLICANT: Bryan, Bruce
APPLICANT: Szent-Gyorgyi, Christopher
TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE
TITLE OF INVENTION: AND FLUORESCENT PROTEINS AND THE USE THEREOF IN DIAGNOSTICS,
TITLE OF INVENTION: SCREENING AND NOVELTY ITEMS
FILE REFERENCE: 24729-121B
CURRENT APPLICATION NUMBER: US/09/609,161B
CURRENT FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 09/277,716
PRIOR FILING DATE: 1999-03-26
PRIOR APPLICATION NUMBER: 60/102,939
PRIOR FILING DATE: 1998-10-01
PRIOR APPLICATION NUMBER: 60/089,367
PRIOR FILING DATE: 1998-06-15
PRIOR APPLICATION NUMBER: 60/079,624
PRIOR FILING DATE: 1998-03-27
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 958
TYPE: DNA
ORGANISM: Aequorea (luminescent jellyfish)
FEATURE:
NAME/KEY: CDS

LOCATION: (115)..(702)
OTHER INFORMATION: Apoaequorin-encoding gene
PUBLICATION INFORMATION:
PATENT DOCUMENT NUMBER: US 07/105,602
PATENT FILING DATE: 1987-10-08
PUBLICATION DATE: 1992-03-03
PUBLICATION INFORMATION:
AUTHORS: Inouye, S.
AUTHORS: No. 643682uchi, M.
AUTHORS: Sakagi, Y.
AUTHORS: Takagi, Y.
AUTHORS: Miyata, T.
AUTHORS: Iwanaga, S.
AUTHORS: Miyata, T.
AUTHORS: Tsuji, F.I.
TITLE: Cloning and sequence analysis of cDNA for the luminescent protein
TITLE: aequorin
JOURNAL: Proc. Natl. Acad. Sci. USA
VOLUME: 82(10)
PAGES: 3154-3158
DATE: 1985-05
US-09-609-161B-5

Query Match 11.9%; Score 38.8; DB 4; Length 958;
Best Local Similarity 58.8%; Pred. No. 0.16;
Matches 67; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 212 GTGATAGCTCTTTGCTTTGGCAAACTTACTTAATCCAAAGGGTTTCTCTGTACAG 271
DB 812 GTGTGATTTTGTATTTAGGACGATTAATCGAATGATTTAGTTTATTAACCA 871

QY 272 TAGATTTCCAAATGTAATAACTTTAATATAAGTAATAAAAAAAAAAAAAA 325
DB 872 CAGAACTTACAAATCGAAAAAGTAATAAAAAAAAAAAAAAAAAAAAAA 925

RESULT 14
US-08-990-103-5
Sequence 5, Application US/08990103
Patent No. 6458547
GENERAL INFORMATION:
APPLICANT: Bruce J. Bryan
APPLICANT: Stephen Gaalema
APPLICANT: Randall B. Murphy
TITLE OF INVENTION: APPARATUS AND METHOD FOR DETECTING AND
TITLE OF INVENTION: IDENTIFYING INSECTICIOUS AGENTS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: CA
COUNTRY: USA
ZIP: 92101-2926
COMPUTER READABLE FORM:
MEDIUM TYPE:
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/990,103
FILING DATE: 12-DEC-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/037,675,
FILING DATE: 02-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/033,745
FILING DATE: 12-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6680-112

TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-238-0999
TELEFAX: 619-238-0062
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 115..702
OTHER INFORMATION: apoaequorin-encoding gene
PUBLICATION INFORMATION:
AUTHORS: Inouye et al.
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
VOLUME: 82
PAGES: 3154-3158
DATE: (1985)
DOCUMENT NUMBER: 5,093,240
US-08-990-103-5

Query Match 11.9%; Score 38.8; DB 4; Length 958;
Best Local Similarity 58.8%; Pred. No. 0.16;
Matches 67; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 212 GTGATAGCTCTTTGCTTTGGCAAACTTACTTAATCCAAAGGGTTTCTCTGTACAG 271
DB 812 GTGTGATTTTGTATTTAGGACGATTAATCGAATGATTTAGTTTATTAACCA 871

QY 272 TAGATTTCCAAATGTAATAACTTTAATATAAGTAATAAAAAAAAAAAAAA 325
DB 872 CAGAACTTACAAATCGAAAAAGTAATAAAAAAAAAAAAAAAAAAAAAA 925

RESULT 15
US-08-446-855A-1
Sequence 1, Application US/08446855A
Patent No. 5849573
GENERAL INFORMATION:
APPLICANT: Stewart, Thomas S
APPLICANT: O'Sullivan, William J
APPLICANT: Flores, Maria V
TITLE OF INVENTION: Nucleotide sequence encoding carbamoyl
TITLE OF INVENTION: phosphate synthetase II
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon & Vanderbye PC
STREET: 1100 No. 584573th Glebe Road, 8th Floor
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,855A
FILING DATE: 06-Jul-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mitchard, Leonard C
REGISTRATION NUMBER: 29,009
REFERENCE/DOCKET NUMBER: 47-80
TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-816-4000
 TELFAX: 703-816-4100
 INFORMATION FOR SRQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 8920 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: genomic
 US-08-446-855A-1

Query Match 11.8%; Score 38.4; DB 2; Length 8920;
 Best Local Similarity 56.2%; Pred. No. 0.48;
 Matches 72; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 198 ATATTTTCCAAAGATGATGCTTTTGGCAAACTCTACTTATCCAAATGGGT 257
 Db 8631 AATTTTGTATATATACAAATTTATTATTCACTCATATGATMAACCAAAATGGTT 8690
 QY 258 TTTTCTGTACAGTAGATTTTCCAAATGTAATAACTTATATAAGTAAAGTAAAAA 317
 Db 8691 TTTTCATTTCAAAATTAATTTTATAATTTAATAATTTAATTAATTAATTAATTA 8750
 QY 318 AAAAAAA 325
 Db 8751 AATATATA 8758

Search completed: January 10, 2003, 03:22:58
 Job time : 58.99 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compen Ltd.

OM nucleic - nucleic search, using SW model

Run on: January 10, 2003, 02:45:21 / Search time 74.3322 Seconds

(Without alignments)
1924.247 Million cell updates/sec

Title: US-09-825-682a-57

Perfect score: 325

Sequence: 1 aaagaagggcagggcgcct.....gtcaaaaaaaaaaaaaa 325

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 389086 seqs, 220051671 residues

Total number of hits satisfying chosen parameters: 778172

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	291.4	89.7	1763	10	US-09-925-302-8
2	285.8	87.9	4797	10	US-09-735-705-134
3	285.8	87.9	4797	10	US-09-850-716A-134
4	285.8	87.9	4797	10	US-09-880-107-3339
5	285.8	87.9	4797	10	US-09-897-778-134
6	144.4	44.4	492	9	US-09-736-457-1041
7	144.4	44.4	492	9	US-09-902-941-1041
8	144.4	44.4	492	9	US-09-849-626-1041
9	120.4	37.0	253	9	US-10-079-623-76
10	44.6	13.7	748	10	US-09-910-943-361
11	42.8	13.2	617	10	US-09-764-877-121
12	41.6	12.8	375	10	US-09-960-352-13618
13	39.8	12.2	659158	9	US-09-771-208-20
14	39.6	12.2	277	10	US-09-960-352-12673
15	39.4	12.1	241	10	US-09-960-352-7904
16	38.8	11.9	289	10	US-09-880-107-1117
17	38.8	11.9	958	10	US-09-803-211-5
18	38.8	11.9	958	10	US-09-746-485A-5
19	38.6	11.9	215	10	US-09-960-352-5093

20	38.6	11.9	2455	10	US-09-918-909-25	Sequence 25, Appl
21	38.6	11.9	2509	10	US-09-925-301-540	Sequence 540, Appl
22	38.4	11.8	325	10	US-09-764-846-30	Sequence 30, Appl
23	38.4	11.8	621	10	US-09-764-846-105	Sequence 105, Appl
24	38.4	11.8	2000	9	US-09-938-842A-3307	Sequence 3307, Ap
25	38.2	11.8	1319	10	US-09-969-347-179	Sequence 179, App
26	38	11.7	819	9	US-10-202-193-236	Sequence 236, App
27	38	11.7	3716	9	US-09-978-695A-210	Sequence 210, App
28	38	11.7	3716	9	US-09-978-697-210	Sequence 210, App
29	38	11.7	3716	9	US-09-978-832A-210	Sequence 210, App
30	38	11.7	3716	9	US-09-999-832A-210	Sequence 210, App
31	38	11.7	3716	9	US-09-978-189-210	Sequence 210, App
32	37.8	11.6	401	10	US-09-960-352-10503	Sequence 10503, A
33	37.8	11.6	1300	10	US-09-822-849A-3	Sequence 3, Appl1
34	37.6	11.6	291	10	US-09-960-352-1243	Sequence 1243, Ap
35	37.6	11.6	368	10	US-09-834-978-47	Sequence 47, Appl
36	37.4	11.5	102	10	US-09-998-598-1903	Sequence 1903, Ap
37	37.4	11.5	444	10	US-09-960-352-1281	Sequence 1281, Ap
38	37.4	11.5	1409	10	US-09-925-301-176	Sequence 176, Appl
39	37.4	11.5	1422	9	US-09-798-889-24	Sequence 24, Appl
40	37.4	11.5	2660	10	US-09-925-299-80	Sequence 80, Appl
41	37.2	11.4	2103	10	US-09-883-060-1	Sequence 1, Appl1
42	37	11.4	469	10	US-09-954-456-248	Sequence 248, App
43	37	11.4	469	10	US-09-954-456-476	Sequence 476, App
44	37	11.4	3110	10	US-09-764-877-3912	Sequence 3912, Ap
45	37	11.4	3110	10	US-09-764-877-3914	Sequence 3914, Ap

ALIGNMENTS

RESULT 1
US-09-925-302-8
Sequence 8, Application US/09925302
Patent No. US20020044941A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: P1104
CURRENT APPLICATION NUMBER: US/09/925.302
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05918
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124.270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 896
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 1763
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-302-8

Query Match 89.7%; Score 291.4; DB 10; Length 1763;

Best Local Similarity 99.3%; Pred. No. 3.2e-60; Matches 303; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY	4	GAGGCGGCGAGGCGCTGAGATCTCTCGACACAGCGCCCTCTGCTGCGCGG	63
DB	752	GAGGCGGCGAGGCGCTGAGATCTCTCGACACAGCGCCCTCTGCTGCGCGG	811
QY	64	TCTCCAGGGGCTGCTTCTCTCTGGAATTGACGAGGGGTCTCTTGGGCGAGCTGGCTCT	123
DB	812	TCTCCAGGGGCTGCTTCTCTCTGGAATTGACGAGGGGTCTCTTGGGCGAGCTGGCTCT	871
QY	124	GAGCGCTTCATCCAGGCGAGGATCTCCGTACTCTCTGAGCCCACTCGGCGCTTG	183
DB	872	GAGCGCTTCATCCAGGCGAGGATCTCCGTACTCTCTGAGCCCACTCGGCGCTTG	930
QY	184	GGCTGGAATCAGGAATATTTTCAAGAAGTATGCTTTTGGTTTGGCAAACTTAC	243
DB	931	GGCTGGAATCAGGAATATTTTCAAGAAGTATGCTTTTGGTTTGGCAAACTTAC	990

Qy	244	TTAATCCAAAGGGTTTTCTGTACAGAGAGATTTCCAAATGATATAAATTAAATATA	303
Db	991	TTAAACCAAGGTTTTCTGTACAGAGATTTCCAAATGATATAAATTAAATATA	1050
Qy	304	AAGTA	308
Db	1051	AAGTA	1055

RESULT 2

```

US-09-735-705-134
Sequence 134, Application US/09735705
Patent No. US20020052329A1
GENERAL INFORMATION:
APPLICANT: Mang, Tongtong
APPLICANT: Fan, Liguu
APPLICANT: Kalos, Michael D.
APPLICANT: Bangur, Chantanya S.
APPLICANT: Hosken, Nancy
APPLICANT: Fanger, Gary R.
APPLICANT: Li, Samuel X.
APPLICANT: Mang, Aijun
APPLICANT: Skeky, Yasir A.W.
APPLICANT: Henderson, Robert A.
APPLICANT: McNeill, Patricia D.
APPLICANT: Fanger, Neil
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.455C14
CURRENT APPLICATION NUMBER: US/09/735,705
CURRENT FILING DATE: 2000-12-12
NUMBER OF SEQ ID NOS: 419
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 134
LENGTH: 4797
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(4797)
OTHER INFORMATION: n = A,T,C or G
US-09-735-705-134

```

Query Match	87.9%;	Score 285.8;	DB 10;	Length 4797;
Best Local Similarity	96.1%;	Pred. No. 9.7e-59;		
Matches 293; Conservative	0;	Mismatches 12;	Indels 0;	Gaps 0;

RESULT 4

```

US-03-880-107-3939
? Sequence 3939, Application US/09880107
? Patent No. US/20020142981A1
? GENERAL INFORMATION:
? APPLICANT: Horne, Darci T.
? APPLICANT: Vockley, Joseph G.
? APPLICANT: Scherf, Uwe
? APPLICANT: Gene Logic, Inc.
? TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
? PIR REFERENCE: 44921-5028-WO
? CURRENT APPLICATION NUMBER: US/09/880,107
? CURRENT FILING DATE: 2001-06-14
? PRIOR APPLICATION NUMBER: US 60/211,379
? PRIOR FILING DATE: 2000-06-14
? PRIOR APPLICATION NUMBER: US 60/237,054
? PRIOR FILING DATE: 2000-10-02
? NUMBER OF SEQ ID NOS: 3950
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 3939

```

```

; LENGTH: 4797
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 Z48199
; NAME/KEY: unsure
; LOCATION: (1)...(4797)
; OTHER INFORMATION: n = a or c or g or t
US-09-880-107-3939

Query Match      87.9%; Score 285.8; DB 10; Length 4797;
Best Local Similarity 96.1%; Pred. No. 9.7e-59;
Matches 293; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 4 GAGGGGCGAGGGGCTGAGATCCTCTGACAGACGCCCGCTGCTGTGGGCGG 63
Db 3765 GAGGGGCGAGGGGCTGAGATCCTCTGACAGGCTCAGGCCGCTCTCTGTGGGCGG 3824
Qy 64 TCTCCAGGGGCTGCTCTCTCTGGAATTTGACGAGGGGTGTCTTGGGCAAGCTGGCTCT 123
Db 3825 TCTCCAGGGGCTGCTCTCTCTGGAATTTGACGAGGGGTGTCTTGGGCAAGCTGGCTCT 3884
Qy 124 GAGCGCTTCATCCAGGCCAGGTTCTCCGTTAGCTCTGTGGGCCCACTGGGCGCTG 183
Db 3885 GAGCGCTTCATCCAGGCCAGGTTCTCCGTTAGCTCTGTGGGCCCACTGGGCGCTG 3944
Qy 184 GGCTGGAATCAGGAATATTTTCCAAAGAGTGAATCTTTTGGGCAAACTCTAC 243
Db 3945 GGCTGGAATCAGGAATATTTTCCAAAGAGTGAATCTTTTGGGCAAACTCTAC 4004
Qy 244 TTATCCAAATGGGTTTTCTCTGTACAGTAGATTTTCCAAATGTATTAATCTTAATA 303
Db 4005 TTATCCAAATGGGTTTTCTCTGTACAGTAGATTTTCCAAATGTATTAATCTTAATA 4064
Qy 304 AAGTA 308
Db 4065 AAGTA 4069

RESULT 5
US-09-897-778-134
; Sequence 134, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Matherakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvik, Thomas S.
; APPLICANT: Carter, Darick
; APPLICANT: Metanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 134
; LENGTH: 4797
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 135_501, 4421, 4467, 4468, 4658
; OTHER INFORMATION: n = A,T,C or G
US-09-897-778-134

Query Match      87.9%; Score 285.8; DB 10; Length 4797;
Best Local Similarity 96.1%; Pred. No. 9.7e-59;
Matches 293; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
```

```

Qy 4 GAGGGGCGAGGGGCTGAGATCCTCTGACAGACGCCCGCTGCTGTGGGCGG 63
Db 3765 GAGGGGCGAGGGGCTGAGATCCTCTGACAGGCTCAGGCCGCTCTCTGTGGGCGG 3824
Qy 64 TCTCCAGGGGCTGCTCTCTCTGGAATTTGACGAGGGGTGTCTTGGGCAAGCTGGCTCT 123
Db 3825 TCTCCAGGGGCTGCTCTCTCTGGAATTTGACGAGGGGTGTCTTGGGCAAGCTGGCTCT 3884
Qy 124 GAGCGCTTCATCCAGGCCAGGTTCTCCGTTAGCTCTGTGGGCCCACTGGGCGCTG 183
Db 3885 GAGCGCTTCATCCAGGCCAGGTTCTCCGTTAGCTCTGTGGGCCCACTGGGCGCTG 3944
Qy 184 GGCTGGAATCAGGAATATTTTCCAAAGAGTGAATCTTTTGGGCAAACTCTAC 243
Db 3945 GGCTGGAATCAGGAATATTTTCCAAAGAGTGAATCTTTTGGGCAAACTCTAC 4004
Qy 244 TTATCCAAATGGGTTTTCTCTGTACAGTAGATTTTCCAAATGTATTAATCTTAATA 303
Db 4005 TTATCCAAATGGGTTTTCTCTGTACAGTAGATTTTCCAAATGTATTAATCTTAATA 4064
Qy 304 AAGTA 308
Db 4065 AAGTA 4069

RESULT 6
US-09-736-457-1041/C
; Sequence 1041, Application US/09736457
; Patent No. US20020168637A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvik, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Manion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1041
; LENGTH: 492
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(492)
; OTHER INFORMATION: n = A,T,C or G
US-09-736-457-1041

Query Match      44.4%; Score 144.4; DB 9; Length 492;
Best Local Similarity 98.6%; Pred. No. 1.5e-25;
Matches 145; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 162 TGTGCCCCACCTGCGGCTGAGTGAATATTTTCCAAAGAGTAGTAGTCT 221
Db 492 TGTGCCCCACCTGCGGCTGAGTGAATATTTTCCAAAGAGTAGTAGTCT 433
Qy 222 TTTGCTTTTGGCAAACTCTACTTAATCCATGGGTTTTTCTCTGTACAGTAGATTTTCC 281
Db 432 TTTGCTTTTGGCAAACTCTACTTAATCCATGGGTTTTTCTCTGTACAGTAGATTTTCC 373
Qy 282 AATGTAAATTAATCTTAATTAAGTA 308
Db 372 AATGTAAATTAATCTTAATTAAGTA 346
```

RESULT 7

US-09-902-941-1041/C

Sequence 1041, Application US/09902941

Patent No. US2002017952A1

GENERAL INFORMATION:

APPLICANT: Henderson, Robert A.

APPLICANT: Wang, Tongtong

APPLICANT: Watanabe, Yoshihiro

APPLICANT: Johnson, Jeffrey C.

APPLICANT: Retter, Marc W.

APPLICANT: Marnetakis, Margarita

APPLICANT: Carter, Darick

APPLICANT: Ranger, Gary R.

APPLICANT: Vedavick, Thomas S.

APPLICANT: Bangur, Chaitanya S.

APPLICANT: McNabb, Andria

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

FILE REFERENCE: 210121.478C17

CURRENT APPLICATION NUMBER: US/09/902,941

CURRENT FILING DATE: 2001-07-10

NUMBER OF SEQ ID NOS: 2002

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1041

LENGTH: 492

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc_feature

LOCATION: 473

OTHER INFORMATION: n = A,T,C or G

US-09-902-941-1041

Query Match 44.4%; Score 144.4; DB 9; Length 492;

Best Local Similarity 98.6%; Pred. No. 1.5e-25;

Matches 145; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 162 TGTGGCCCACTGGGCGCTGGGCGGATCGAATATTTTCCAAAGATGATGCT 221

DB 492 TGTGGCCCACTGGGCGCTGGGCGGATCGAATATTTTCCAAAGATGATGCT 433

QY 222 TTTCCTTTGGCAAACTCTACTTAATCAATGAGTTTCTCTGACAGTAGATTTTCC 281

DB 432 TTTCCTTTGGCAAACTCTACTTAATCAATGAGTTTCTCTGACAGTAGATTTTCC 373

QY 282 AAATGTAATACTTAATATAAAGTA 308

DB 372 AAATGTAATACTTAATATAAAGTA 346

RESULT 8

US-09-849-626-1041/C

Sequence 1041, Application US/09849626

Patent No. US20020197669A1

GENERAL INFORMATION:

APPLICANT: Bangur, Chaitanya

APPLICANT: Ranger, Gary

APPLICANT: Wang, Tongtong

APPLICANT: Wang, Tongtong

APPLICANT: Switzer, Anne

APPLICANT: McNeill, Patricia

APPLICANT: Clapper, Jonathan

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER

FILE REFERENCE: 210121.478C16

CURRENT APPLICATION NUMBER: US/09/849,626

CURRENT FILING DATE: 2001-05-03

NUMBER OF SEQ ID NOS: 1926

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 1041

LENGTH: 492

TYPE: DNA

ORGANISM: Homo sapien

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1)...(492)

OTHER INFORMATION: n = A,T,C or G

US-09-849-626-1041

Query Match 44.4%; Score 144.4; DB 9; Length 492;

Best Local Similarity 98.6%; Pred. No. 1.5e-25;

Matches 145; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 162 TGTGGCCCACTGGGCGCTGGGCGGATCGAATATTTTCCAAAGATGATGCT 221

DB 492 TGTGGCCCACTGGGCGCTGGGCGGATCGAATATTTTCCAAAGATGATGCT 433

QY 222 TTTCCTTTGGCAAACTCTACTTAATCAATGAGTTTCTCTGACAGTAGATTTTCC 281

DB 432 TTTCCTTTGGCAAACTCTACTTAATCAATGAGTTTCTCTGACAGTAGATTTTCC 373

QY 282 AAATGTAATACTTAATATAAAGTA 308

DB 372 AAATGTAATACTTAATATAAAGTA 346

RESULT 9

US-10-079-623-76

Sequence 76, Application US/10079623

Patent No. US20020169302A1

GENERAL INFORMATION:

APPLICANT: Havukkala, Ilkka J.

APPLICANT: Glenn, Matthew

APPLICANT: Gligor, Murray R.

APPLICANT: Molenaar, Adrian J.

TITLE OF INVENTION: Compositions isolated from bovine

TITLE OF INVENTION: mammary gland and methods for their use.

FILE REFERENCE: 11000.1044C3

CURRENT APPLICATION NUMBER: US/10/079,623

CURRENT FILING DATE: 2002-02-19

NUMBER OF SEQ ID NOS: 370

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 76

LENGTH: 253

TYPE: DNA

ORGANISM: Bovine

US-10-079-623-76

Query Match 37.0%; Score 120.4; DB 9; Length 253;

Best Local Similarity 75.6%; Pred. No. 5.8e-20;

Matches 189; Conservative 0; Mismatches 56; Indels 5; Gaps 3;

QY 77 CTTCCTCTGAAATTGACGAGGCGGTCTTGGGACAGCTGGCTGAGGCGCTGCATC 136

DB 4 CTTCCTCTGTAAGTACAGACAGCTCTTGGGTACCTGACCTGAGTGCCACG 63

QY 137 CAAGCCAGGTTCTCCCTTACCTCTGAGCCCACTGGGCGCTGGATGAG 196

DB 64 GATGACAGGCTCACTTTGATAGCT--CTGTAGCCTGACTGTGGCCGAGATCAG 121

QY 197 AATATTTTCCAAAGAGTATGCTTTTGGCTTTGGCAAAAC-TCCTATTATCAATGG 255

DB 122 AATA--TTCCAAAGAGTATGCTTTTGGCTTTGGCCAACTTTATTATTAATCAATGG 179

QY 256 GTTTTCTCTGACAGTGAATTTCCAAATGATACTTAATATAAGTAAAAA 315

DB 180 GTTTTCTCTGACAGTGAATTTTCCAAATGATACTTAATATAAGTATGCT 239

QY 316 AAAAAAAAA 325

DB 240 GAAAAAAAA 249

RESULT 10

```

US-09-910-943-361
? Sequence 361, Application US/09910943
? Patent No. US20020081610A1
? GENERAL INFORMATION:
? APPLICANT: Hemmati-Bitvanlou, Ali
? APPLICANT: Altman, Curtis
? TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
? FILE REFERENCE: 7529/1G148US1
? CURRENT APPLICATION NUMBER: US/09/910,943
? CURRENT FILING DATE: 2001-07-23
? NUMBER OF SEQ ID NOS: 742
? SOFTWARE: PatentIn version 3.1
? SEQ ID NO 361
? LENGTH: 748
? TYPE: DNA
? ORGANISM: Xenopus laevis
? FEATURE:
? NAME/KEY: misc feature
? LOCATION: (1) ..(748)
? OTHER INFORMATION: n may be a o r c or c/u
US-09-910-943-361

```

```

/ LOCATION: (546998)..(547017)
/ OTHER INFORMATION: n is unidentified a, c, g, or t
/ NAME/KEY: misc feature
/ LOCATION: (494715)..(494814)
/ OTHER INFORMATION: n is unidentified a, c, g, or t
/ NAME/KEY: misc feature
/ LOCATION: (390986)..(391005)
/ OTHER INFORMATION: n is unidentified a, c, g, or t
/ NAME/KEY: misc feature
/ LOCATION: (346860)..(346823)
/ OTHER INFORMATION: n is unidentified a, c, g, or t
/ NAME/KEY: misc feature
/ LOCATION: (317174)..(317193)
/ OTHER INFORMATION: n is unidentified a, c, g, or t
/ NAME/KEY: misc feature
/ LOCATION: (280353)..(280373)
/ OTHER INFORMATION: n is unidentified a, c, g, or t
/ NAME/KEY: misc feature
/ LOCATION: (271829)..(271848)
/ OTHER INFORMATION: n is unidentified a, c, g, or t
/ NAME/KEY: misc feature
/ LOCATION: (183872)..(183891)
/ OTHER INFORMATION: n is unidentified a, c, g, or t
/ NAME/KEY: misc feature
/ LOCATION: (170625)..(170645)
/ OTHER INFORMATION: n is unidentified a, c, g, or t
/ NAME/KEY: misc feature
/ LOCATION: (132680)..(132700)
/ OTHER INFORMATION: n is unidentified a, c, g, or t
/ NAME/KEY: misc feature
/ OTHER INFORMATION: n is a, c, g, or t
US-09-771-208-20

```

```

Query Match 12.2%; Score 39.8; DB 9; Length 659158;
Best Local Similarity 47.1%; Pred. No. 11;
Matches 122; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

```

```

QY 63 GTCCTCCAGGGGCTCTCTCTCTGGAATTGACGAGGGGTCTTGGCAGAGTGGCTC 122
Db 39253 GTTTTGAAGAGGTGTATATGAGACAGCGCTCCAGGGGTGGGGGTTAAAGCTG 39312

QY 123 TGAGCGCTCCATCCAGGCGCAGGTTCCGTTAGCTCTGGGGCCCACTGGGCGCT 182
Db 39313 TGAAGTGAAGCTTGGCAAGTGGCTCCGTAAGTAAAGTCTTCTCTAAGACTATGACT 39372

QY 183 GGGCTGGATCAGCAATATTTTCCAAAGATGATAGTCTTTTGGCAAACTCTA 242
Db 39373 GAGTTGATCCCGAATCCATGTGTGSAAGAGAGACTACTGTCTCTGGCCCCCA 39432

QY 243 CTATATCAATGGGTTTCTCTGTACATGATTTTCCAAATGTAATAACTTAAAT 302
Db 39433 CATGGGACCAAGGCTACTGAGTACAGCCCATGTCCTCAGAAAGAAAGTCAATG 39492

QY 303 AAAGTAAAAAAGAAAAA 321
Db 39493 AATTAAAAAAGAAAAA 39511

```

```

RESULT 14
US-09-960-352-12673/C
/ Sequence 12673, Application US/09960352
/ Patent No. US20020137139A1
/ GENERAL INFORMATION:
/ APPLICANT: Warren, Wesley C.
/ APPLICANT: Tao, Nengbing
/ APPLICANT: Byatt, John C.
/ APPLICANT: Mathalagan, Nagappan
/ TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
/ FILE REFERENCE: 16511.006/37-21(10298)C
/ CURRENT APPLICATION NUMBER: US/09/960,352
/ CURRENT FILING DATE: 2001-09-24
/ NUMBER OF SEQ ID NOS: 15112

```

```

/ SEQ ID NO 12673
/ LENGTH: 277
/ TYPE: DNA
/ ORGANISM: Bos taurus
/ OTHER INFORMATION: Clone ID: 54-LIB3058-039-Q1-K1-P10
US-09-960-352-12673

```

```

Query Match 12.2%; Score 39.6; DB 10; Length 277;
Best Local Similarity 56.0%; Pred. No. 0.86;
Matches 75; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

```

```

QY 192 TCAGAAATTTTCCAAAGATGATGTTGCTTTGGCAAACTCTACTATCCA 251
Db 219 TCAGAAATTTTTCATTTTAAAAATTTTCTTTTAAATTTTAAATTTT 160

QY 252 ATGGGTTTCTCTGTACATGATTTTCCAAATGTAATAACTTAAATAAAGTAAA 311
Db 159 TTTTCTTTTCTTTTAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTT 100

QY 312 AAAAAAAAAAAAAA 325
Db 99 AAAAAAAAAAAAAA 86

```

```

RESULT 15
US-09-960-352-7904/C
/ Sequence 7904, Application US/09960352
/ Patent No. US20020137139A1
/ GENERAL INFORMATION:
/ APPLICANT: Warren, Wesley C.
/ APPLICANT: Tao, Nengbing
/ APPLICANT: Byatt, John C.
/ APPLICANT: Mathalagan, Nagappan
/ TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
/ FILE REFERENCE: 16511.006/37-21(10298)C
/ CURRENT APPLICATION NUMBER: US/09/960,352
/ CURRENT FILING DATE: 2001-09-24
/ NUMBER OF SEQ ID NOS: 15112
/ SEQ ID NO 7904
/ LENGTH: 241
/ TYPE: DNA
/ ORGANISM: Bos taurus
/ OTHER INFORMATION: Clone ID: 34-LIB3057-019-Q1-K1-A10
US-09-960-352-7904

```

```

Query Match 12.1%; Score 39.4; DB 10; Length 241;
Best Local Similarity 67.9%; Pred. No. 0.91;
Matches 55; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

```

```

QY 245 TAATCAATGGGTTTCTCTGTACATGATTTTCCAAATGTAATAACTTAAATTA 304
Db 211 TAATGAATTAATGAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAA 152

QY 305 AGTAAAAAAGAAAAA 325
Db 151 AATAAAAAAGAAAAA 131

```

```

Search completed: January 10, 2003, 04:50:01
Job time : 196.312 secs

```


/clone="IMAGE:1858507"
 /clone_lib="Soares NtHMPu S1"
 /tissue_type="Pooled human melanocyte, fetal heart, and
 pregnant uterus"
 /lab_host="DH10B"
 /note="Organ: mixed (see below); Vector: pT7T3D-Pac
 (Pharmacia) with a modified polylinker; Site 1: Not 1;
 Site 2: Eco RI. Equal amounts of plasmid DNA from three
 normalized libraries (melanocyte 2NBH, pregnant uterus
 NBHPU, and fetal heart NBH19H) were mixed, and ss circles
 were made in vitro. Following HAP purification, this DNA
 was used as tracer in a subtractive hybridization
 reaction. The driver was PCR-amplified cDNAs from pools of
 5,000 clones made from the same 3 libraries. The pools
 consisted of I.M.A.G.E. clones 260232-265223,
 340488-345479, and 484488-489479."

BASE COUNT 117 a 146 c 137 g 110 t 3 others
 ORIGIN

Query Match 98.6%; Score 320.4; DB 9; Length 513;
 Best Local Similarity 99.7%; Pred. No. 2.4e-42;
 Matches 321; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GAGGGGGGAGGGGCTCGAGATCTCTCGACACACGCCCTGCTGCTGGGCGCG 63
 DB 323 GAGGGGGGAGGGGCTCGAGATCTCTCGACACACGCCCTGCTGCTGGGCGCG 264

QY 64 TCTCCAGGGGCTGCTCTCTCTGAAATTGACGAGGGGTGTCTTGGGAGAGCTGCT 123
 DB 263 TCTCCAGGGGCTGCTCTCTCTGAAATTGACGAGGGGTGTCTTGGGAGAGCTGCT 204

QY 124 GAGCGCTTCATCCAGGCCAGGCTTCCTCGTAGTCTCTGTGGCCACCTGGGCGCTG 183
 DB 203 GAGCGCTTCATCCAGGCCAGGCTTCCTCGTAGTCTCTGTGGCCACCTGGGCGCTG 144

QY 184 GCGTGAATCAGGAATTTTCCAAAGATGATCTTTGCTTTGGCAAACCTAC 243
 DB 143 GCGTGAATCAGGAATTTTCCAAAGATGATCTTTGCTTTGGCAAACCTAC 84

QY 244 TTAATCAATGGGTTTCTCTGTACAGTAGATTTTCCAAATGTAATACTTAATA 303
 DB 83 TTAATCAATGGGTTTCTCTGTACAGTAGATTTTCCAAATGTAATACTTAATA 24

QY 304 AAGTAAAAAAAAAAAAA 325
 DB 23 AAGTAAAAAAAAAAAAA 2

RESULT 2
 B0644995 534 bp mRNA linear EST 15-JUL-2002
 LOCUS B0644995
 DEFINITION AGENCOURT_8498888 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6297772
 5', mRNA sequence.
 ACCESSION B0644995
 VERSION B0644995.1 GI:21769167
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 534)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: rgs@bbs-remail.nih.gov
 Tissue Procurement: CGAP (Stanford)
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov

Plate: L1CM2507 row: m column: 05
 High quality sequence stop: 532.
 Location/Qualifiers
 source
 1. 534
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6297772"
 /clone_lib="NIH_MGC_100"
 /tissue_type="hepatocellular carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: liver; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGACAGAG(G). Size-selected >500bp for average insert size
 1.8kb. Library constructed by Ling Hong in the laboratory
 of Gerald M. Rubin (University of California, Berkeley)
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC
 library."

BASE COUNT 110 a 137 c 153 g 126 t 8 others
 ORIGIN

Query Match 98.6%; Score 320.4; DB 14; Length 534;
 Best Local Similarity 99.7%; Pred. No. 2.4e-42;
 Matches 321; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GAGGGGGGAGGGGCTCGAGATCTCTCGACACACGCCCTGCTGCTGGGCGCG 63
 DB 203 GAGGGGGGAGGGGCTCGAGATCTCTCGACACACGCCCTGCTGCTGGGCGCG 262

QY 64 TCTCCAGGGGCTGCTCTCTCTGAAATTGACGAGGGGTGTCTTGGGAGAGCTGCT 123
 DB 263 TCTCCAGGGGCTGCTCTCTCTGAAATTGACGAGGGGTGTCTTGGGAGAGCTGCT 322

QY 124 GAGCGCTTCATCCAGGCCAGGCTTCCTCGTAGTCTCTGTGGCCACCTGGGCGCTG 183
 DB 323 GAGCGCTTCATCCAGGCCAGGCTTCCTCGTAGTCTCTGTGGCCACCTGGGCGCTG 382

QY 184 GCGTGAATCAGGAATTTTCCAAAGATGATCTTTGCTTTGGCAAACCTAC 243
 DB 383 GCGTGAATCAGGAATTTTCCAAAGATGATCTTTGCTTTGGCAAACCTAC 442

QY 244 TTAATCAATGGGTTTCTCTGTACAGTAGATTTTCCAAATGTAATACTTAATA 303
 DB 443 TTAATCAATGGGTTTCTCTGTACAGTAGATTTTCCAAATGTAATACTTAATA 502

QY 304 AAGTAAAAAAAAAAAAA 325
 DB 503 AAGTAAAAAAAAAAAAA 524

RESULT 3
 B0644995 722 bp mRNA linear EST 21-MAR-2002
 LOCUS B0644995/c
 DEFINITION UI-CF-BN1-add-c-24-0-UI s1 UI-CF-BN1 Homo sapiens cDNA clone
 UI-CF-BN1-add-c-24-0-UI 3', mRNA sequence.
 ACCESSION B0644995
 VERSION B0644995.1 GI:19602077
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 722)
 Bonaldo, M.F., Lennon, G. and Soares, M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 97044477
 MEDLINE
 COMMENT
 Contact: McCray, PB
 McCray Lab
 University of Iowa
 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA

Tel: 319 356 4866
 Fax: 319 356 7171
 Email: paul-mccray@uiowa.edu
 Tissue Procurement: Dr. M. J. Welsh, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 The following repetitive elements were found in this cDNA
 sequence: 1-39, >AT-richflow_complexity (matched complement)
 617-637, >AT-richflow_complexity (matched complement)
 Seq primer: M13 FORWARD
 POLYA=yes.

FEATURES

source

Location/Qualifiers
 1..722
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="UI-CF-EN1-add-C-24-0-UI"
 /clone_lib="UI-CF-EN1"
 /tissue_type="Primary Lung Cystic Fibrosis Epithelial
 Cells"
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /note="Organ: Lung; Vector: pRT3-Pac (Pharmacia) with a
 modified polylinker; Site 1: EcoR I; Site 2: Not I;
 UI-CF-EN1 is a normalized cDNA library containing the
 following tissue(s): Primary Lung Cystic Fibrosis
 Epithelial Cells. The library was constructed according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was ligated to an EcoR I adaptor, digested with Not
 I, and cloned directionally into pRT3-Pac vector. The
 oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dT)18 tail. The
 sequence tag for this library is CTCCTCAGGT.
 TAG_LIB=UI-CF-EN1
 TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS
 6hr to LPS 24h
 TAG_SEQ=CTGCTCAGGT"
 BASE COUNT 195 a 201 c 170 g 155 t 1 others
 ORIGIN

Query Match

Best Local Similarity 99.7%; Score 320.4; DB 14; Length 722;
 Pred. No. 2e-42; Mismatches 1; Indels 0; Gaps 0;

Matches 321; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

4 GAGGGGCGAGGGGCTGGAGATCTCTGCGACAGCCCGCTGCTGCGGGCGG 63
 322 GAGGGGCGAGGGGCTGGAGATCTCTGCGACAGCCCGCTGCGGGCGG 263
 64 TCTCCAGGGGCTGCTCTCTCTGGAATTGACGAGGGGCTGCGGAGAGCTGCT 123
 262 TCTCCAGGGGCTGCTCTCTCTGGAATTGACGAGGGGCTGCGGAGAGCTGCT 203
 124 GAGCGCTTCATCCAGGCGAGGTTCTCCGTTAGCTCTGTGGCCCACTGCGGCGCT 183
 202 GAGCGCTTCATCCAGGCGAGGTTCTCCGTTAGCTCTGTGGCCCACTGCGGCGCT 143
 184 GGTGGAATCAGGAATATTTTCCAAAGATGATGCTTTTGGCTTTGGCAAACTTAC 243
 142 GGTGGAATCAGGAATATTTTCCAAAGATGATGCTTTTGGCTTTGGCAAACTTAC 83
 244 TTATTCATGAGGTTTCTCTGACAGTAGATTTTCCAAATGTAATTAATTAATA 303
 82 TTATTCATGAGGTTTCTCTGACAGTAGATTTTCCAAATGTAATTAATTAATA 23
 304 AAGTAAAAAATTTTCTCTGACAGTAGATTTTCCAAATGTAATTAATTAATA 325
 22 AAGTAAAAAATTTTCTCTGACAGTAGATTTTCCAAATGTAATTAATTAATA 1

RESULT 4
 BM982212/c
 LOCUS
 DEFINITION
 UI-CF-EN1-acr-o-03-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone
 UI-CF-EN1-acr-o-03-0-UI 3', mRNA sequence.
 ACCESSION
 BM982212
 VERSION
 BM982212.1 GI:19605484
 EST

SOURCE

ORGANISM

human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

```

Db      322 GAGGCGGCGAGGCGCTGAGAGATCCTCCTGACGACCAAGCCCGTCTGCTGCGCGCG 263
QY      64 TCTCCAGGGGCTGCTTCTCTCTGGAATTTGAAGAGGGGTCTTTGGGCGAGACTGCTCT 123
Db      262 TCTCCAGGGGCTGCTTCTCTCTGGAATTTGAAGAGGGGTCTTTGGGCGAGACTGCTCT 203
QY      124 GAGCGCTCCATCCAGAGCGCAGGTTCTCCGTAGCTCTCTGCGCCCACTCGGCGCGCTG 183
Db      202 GAGCGCTCCATCCAGAGCGCAGGTTCTCCGTAGCTCTCTGCGCCCACTCGGCGCGCTG 143
QY      184 GCGTGAATCAGGAATATTTTCCAAAGAGTATAGTCTTTTGTGTCATTAATTA 243
Db      142 GCGTGAATCAGGAATATTTTCCAAAGAGTATAGTCTTTTGTGTCATTAATTA 83
QY      244 TTAATCCAAATGGGTTTTTCTCTGTCAGTATAGTTTCCAAATGTAACTTAATTA 303
Db      82 TTAATCCAAATGGGTTTTTCTCTGTCAGTATAGTTTCCAAATGTAACTTAATTA 23
QY      304 AAGTAAAAAAGGTTTTTCTCTGTCAGTATAGTTTCCAAATGTAACTTAATTA 325
Db      22 AAGTAAAAAAGGTTTTTCTCTGTCAGTATAGTTTCCAAATGTAACTTAATTA 1

RESULT 5
B0023686/c      327 bp      mRNA      linear      EST 27-MAR-2002
LOCUS      B0023686
DEFINITION      UI-1-BB0-abv-g-06-0-UI.s1 NCI CGAP P14 Homo sapiens cDNA clone
ACCESSION      B0023686
VERSION      B0023686.1 GI:19758965
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE      Tumor Gene Index
JOURNAL      Unpublished (1997)
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Dr. Steven Brown
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 1-38, >AT rich#low_complexity
Seq primer: M13 FORWARD
POLY-A=Yes.

FEATURES
source      Location/Qualifiers
1..327
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-1-BB0-abv-g-06-0-UI"
/clone_lib="NCI CGAP P14"
/tissue_type="Placenta"
/dev_stage="Full Term"
/lab_host="DH10B (Life Technologies)"
/note="Organ: Placenta; Vector: pRTT3-Pac (Pharmacia) with
a modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI CGAP P14 is a cDNA library containing the following
tissue(s): Placenta full term. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed,
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pRTT3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The

```

```

sequence tag for this library is AGGAA.
TAG_LIB=UI-1-BB0
TAG_TISSUE=Placenta human full term
TAG_SEQ=AGGAA"
BASE COUNT      83 a      84 c      80 g      79 t      1 others
ORIGIN
Query Match      98.0%; Score 318.4; DB 14; Length 327;
Best Local Similarity 99.4%; Fred. No. 6.6e-42;
Matches 319; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
4 GAGGCGGCGAGGCGCTGAGAGATCCTCCTGACGACCAAGCCCGTCTGCTGCGCGCG 63
Db      321 GAGGCGGCGAGGCGCTGAGAGATCCTCCTGACGACCAAGCCCGTCTGCTGCGCGCG 262
QY      64 TCTCCAGGGGCTGCTTCTCTCTGGAATTTGAAGAGGGGTCTTTGGGCGAGACTGCTCT 123
Db      261 TCTCCAGGGGCTGCTTCTCTCTGGAATTTGAAGAGGGGTCTTTGGGCGAGACTGCTCT 202
QY      124 GAGCGCTCCATCCAGAGCGCAGGTTCTCCGTAGCTCTCTGCGCCCACTCGGCGCGCTG 183
Db      201 GAGCGCTCCATCCAGAGCGCAGGTTCTCCGTAGCTCTCTGCGCCCACTCGGCGCGCTG 142
QY      184 GCGTGAATCAGGAATATTTTCCAAAGAGTATAGTCTTTTGTGTCATTAATTA 243
Db      141 GCGTGAATCAGGAATATTTTCCAAAGAGTATAGTCTTTTGTGTCATTAATTA 82
QY      244 TTAATCCAAATGGGTTTTTCTCTGTCAGTATAGTTTCCAAATGTAACTTAATTA 303
Db      81 TTAATCCAAATGGGTTTTTCTCTGTCAGTATAGTTTCCAAATGTAACTTAATTA 22
QY      304 AAGTAAAAAAGGTTTTTCTCTGTCAGTATAGTTTCCAAATGTAACTTAATTA 324
Db      21 AAGTAAAAAAGGTTTTTCTCTGTCAGTATAGTTTCCAAATGTAACTTAATTA 1

RESULT 6
B0023676/c      330 bp      mRNA      linear      EST 27-MAR-2002
LOCUS      B0023676
DEFINITION      UI-1-BB0-abv-f-07-0-UI.s1 NCI CGAP P14 Homo sapiens cDNA clone
ACCESSION      B0023676
VERSION      B0023676.1 GI:19758955
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE      Tumor Gene Index
JOURNAL      Unpublished (1997)
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Dr. Steven Brown
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 1-38, >AT rich#low_complexity
Seq primer: M13 FORWARD
POLY-A=Yes.

FEATURES
source      Location/Qualifiers
1..330
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-1-BB0-abv-f-07-0-UI"
/clone_lib="NCI CGAP P14"
/tissue_type="Placenta"
/dev_stage="Full Term"

```

```

/lab_host="DH10B (Life Technologies)"
/notes="Organ: Placenta; Vector: pT73-Pac (Pharmacia) with
a modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI-CGAP Pl4 is a cDNA library containing the following
tissue(s): Placenta full term. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is AGGA.
TAG_LIB=UI-1-B80
TAG_TISSUE=Placenta human full term
TAG_SEQ=AGGA"

BASE COUNT      83 a      85 c      81 g      80 t      1 others
ORIGIN

Query Match      98.0%; Score 318.4; DB 14; Length 330;
Best Local Similarity 99.4%; Pred. No. 6.6e-42;
Matches 319; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      4  GAGGCGGCGAGGGGCTGGAGATCTCTCTGACAGCCAGCCGCTGCGGCGCG 63
DB      321 GAGGCGGCGAGGGGCTGGAGATCTCTCTGACAGCCAGCCGCTGCGGCGCG 262

QY      64  TCTCCAGGGGGCTGCTCTCTCTGAAATTGACGAGGGGTGTCTTGGGCAAGAGCTGCT 123
DB      261 TCTCCAGGGGGCTGCTCTCTCTGAAATTGACGAGGGGTGTCTTGGGCAAGAGCTGCT 202

QY      124 GAGCGCTTCATCCAGGCCAGGTTCCTCGTTAGCTCTGTGGGCCCACTGGGCGCTTG 183
DB      201 GAGCGCTTCATCCAGGCCAGGTTCCTCGTTAGCTCTGTGGGCCCACTGGGCGCTTG 142

QY      184 GGTGGAAATCAGGAATATTTTCCAAAGAGTGAATGATCTTTGCTTTGGGAAAACCTTAC 243
DB      141 GGTGGAAATCAGGAATATTTTCCAAAGAGTGAATGATCTTTGCTTTGGGAAAACCTTAC 82

QY      244 TTAATCCATGGGTTTCTCTGTACAGTAGATTTTCCAAATGTAATAACTTTAATATA 303
DB      81  TTAATCCATGGGTTTCTCTGTACAGTAGATTTTCCAAATGTAATAACTTTAATATA 22

QY      304 AAGTAAAAAAAAAAAAAAAA 324
DB      21  AAGTAAAAAAAAAAAAAAAA 1

RESULT 7
A1831036/c      615 bp      mRNA      linear      EST 21-DEC-1999
LOCUS      wj62a07.x1 NCI CGAP Lu19 Homo sapiens cDNA clone IMAGE:2407380 3'
DEFINITION      similar to gb:J05392 SYNDSCAN-1 PRCUNSOR (HUMAN), contains PMS.13
PMS5 repetitive element ;, mRNA sequence.

ACCESSION      A1831036
VERSION      A1831036.1 GI:5451707
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 615)
AUTHORS      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncigap.
TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL      Unpublished (1997)
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: cgaaps-remail.nih.gov
Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.

```

```

DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/BLIN at:
www.bio.ln1.gov/bdrp/image/image.html
Insert Length: 1533 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence crop: 443.

FEATURES
Source
Location/Qualifiers
1..615
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2407380"
/clone_lib="NCI-CGAP_Lu19"
/tissue_type="squamous cell carcinoma, poorly
differentiated (4 pooled tumors, including primary and
metastatic)"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
pooled lung tumor tissue, and was then primed with a Not I
- oligo(dT) primer. Double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT73 vector. Library went through one round of
normalization. Library constructed by Bento Soares and M.
Fatima Bonaldo."

BASE COUNT      156 a      179 c      153 g      127 t
ORIGIN

Query Match      97.7%; Score 317.4; DB 9; Length 615;
Best Local Similarity 99.7%; Pred. No. 6.6e-42;
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      4  GAGGCGGCGAGGGGCTGGAGATCTCTCTGACAGCCAGCCGCTGCGGCGCG 63
DB      319 GAGGCGGCGAGGGGCTGGAGATCTCTCTGACAGCCAGCCGCTGCGGCGCG 260

QY      64  TCTCCAGGGGGCTGCTCTCTCTGAAATTGACGAGGGGTGTCTTGGGCAAGAGCTGCT 123
DB      259 TCTCCAGGGGGCTGCTCTCTCTGAAATTGACGAGGGGTGTCTTGGGCAAGAGCTGCT 200

QY      124 GAGCGCTTCATCCAGGCCAGGTTCCTCGTTAGCTCTGTGGGCCCACTGGGCGCTTG 183
DB      199 GAGCGCTTCATCCAGGCCAGGTTCCTCGTTAGCTCTGTGGGCCCACTGGGCGCTTG 140

QY      184 GGTGGAAATCAGGAATATTTTCCAAAGAGTGAATGATCTTTGCTTTGGGAAAACCTTAC 243
DB      139 GGTGGAAATCAGGAATATTTTCCAAAGAGTGAATGATCTTTGCTTTGGGAAAACCTTAC 80

QY      244 TTAATCCATGGGTTTCTCTGTACAGTAGATTTTCCAAATGTAATAACTTTAATATA 303
DB      79  TTAATCCATGGGTTTCTCTGTACAGTAGATTTTCCAAATGTAATAACTTTAATATA 20

QY      304 AAGTAAAAAAAAAAAAAAAA 322
DB      19  AAGTAAAAAAAAAAAAAAAA 1

RESULT 8
BQ188527
LOCUS      BQ188527
DEFINITION      UI-E-EJ1-a'w-j-12-0-UI .r1 UI-E-EJ1 Homo sapiens cDNA clone
UI-E-EJ1-a'w-j-12-0-UI 5', mRNA sequence.
ACCESSION      BQ188527
VERSION      BQ188527.1 GI:20364078
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 666)
AUTHORS      Bonaldo,M.F., Lennon,G. and Soares,M.B.

```

TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 9704447
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
The following repetitive elements were found in this cDNA sequence: 627-657, >POLY_AHSimple_repeat
Seq primer: M13 REVERSE.
Location/Qualifiers
1. .666
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-E-EJ1-ajw-j-12-0-UI"
/clone_lib="UI-E-EJ1"
/tissue_type="fetal eyes, lens, eye anterior segment, optic nerve, retina, Retina Foveal and Macular, RPE and Choroid"
/dev_stage="fetal and adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site_1: ECOR I; Site_2: Not I; UI-E-EJ1 is a subtracted cDNA library constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an ECOR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (GT)₁₈ tail. The sequence tags for this library are: fetal eyes, ADAR1CAAGA ; lens, CGATTACGA; eye anterior segment, AATGCCGAT; optic nerve, CCATTAAGT; retina, CCGCG; Retina Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."
BASE COUNT 143 a 160 c 185 g 175 t 3 others
ORIGIN
Query Match 97.5%; Score 316.8; DB 14; Length 666;
Best Local Similarity 98.8%; Pred. No. 7,9e-42;
Matches 318; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 4 GAGGCGGCGAGGGGCTCGAGATCTCTCGACACACGCCCTGCTGCGCGCG 63
DB 328 GAGGCGGCGAGGGGCTCGAGATCTCTCGACACACGCCCTGCTGCGCGCG 387
QY 64 TCTCAGGGGCTGCTTCTCTCGAAATTGACGAGGGGTGCTTGAGCAAGCTGCT 123
DB 388 TCTCAGGGGCTGCTTCTCTCGAAATTGACGAGGGGTGCTTGAGCAAGCTGCT 447
QY 124 GAGGCGCTCCATCCAGAGGCTTCTCGTTAGCTCTGAGCCCACTGAGCCCTG 183
DB 448 GAGGCGCTCCATCCAGAGGCTTCTCGTTAGCTCTGAGCCCACTGAGCCCTG 507
QY 184 GCGTGAATCAGAAATTTTCCAAAGAGTAGTCTTTGCTTTGGCAAACTCTAC 243
DB 508 GCGTGAATCAGAAATTTTCCAAAGAGTAGTCTTTGCTTTGGCAAACTCTAC 567
QY 244 TTATCCAAATGGGTTTTTCTCTGACAGTAGTTTCCAAATGTAATTAATTATA 303

DB 568 TTATCCAAATGGGTTTTTCTCTGACAGTAGTCTCCAAATGTAATTAATTATA 627
QY 304 AAGTAAAAAAAAAAAAAAAA 325
DB 628 AAGTAAAAAAAAAAAAAAAA 649
RESULT 9
AM337157/c 614 bp mRNA linear EST 31-JAN-2000
LOCUS
DEFINITION xw81f06.x1 NCI CGAP Pauli Homo sapiens cDNA clone IMAGE:2834435, similar to gb:J05392 SYNDSCAN-1 PRECURSOR (HUMAN), contains element MER22 repetitive element ;, mRNA sequence.
ACCESSION AM337157
VERSION AM337157.1 GI:6833783
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
TITLE Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/HLN at: www.bio.lnlnl.gov/bdrp/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 359.
Location/Qualifiers
1. .614
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2834435"
/clone_lib="NCI CGAP Pauli"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1,72 kb. Life Technologies catalog #: 11548-013"
BASE COUNT 157 a 178 c 151 g 127 t 1 others
ORIGIN
Query Match 97.2%; Score 315.8; DB 10; Length 614;
Best Local Similarity 99.4%; Pred. No. 1,2e-41;
Matches 317; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4 GAGGCGGCGAGGGGCTCGAGATCTCTCGACACACGCCCTGCTGCGCGCG 63
DB 319 GAGGCGGCGAGGGGCTCGAGATCTCTCGACACACGCCCTGCTGCGCGCG 260
QY 64 TCTCAGGGGCTGCTTCTCTCGAAATTGACGAGGGGTGCTTGAGCAAGCTGCT 123
DB 259 TCTCAGGGGCTGCTTCTCTCGAAATTGACGAGGGGTGCTTGAGCAAGCTGCT 200
QY 124 GAGGCGCTCCATCCAGAGGCTTCTCGTTAGCTCTGAGCCCACTGAGCCCTG 183
DB 199 GAGGCGCTCCATCCAGAGGCTTCTCGTTAGCTCTGAGCCCACTGAGCCCTG 140
QY 184 GCGTGAATCAGAAATTTTCCAAAGAGTAGTCTTTGCTTTGGCAAACTCTAC 243
DB 139 GCGTGAATCAGAAATTTTCCAAAGAGTAGTCTTTGCTTTGGCAAACTCTAC 80
QY 244 TTATCCAAATGGGTTTTTCTCTGACAGTAGTTTCCAAATGTAATTAATTATA 303
DB 79 TTATCCAAATGGGTTTTTCTCTGACAGTAGTTTCCAAATGTAATTAATTATA 20

Db 322 GAGGCGCGCAGGCGCTGAGATCTCTCTGACGACGACGCCGCTGCTGCTGCGCGCG 263

QY 64 TCTCCAGGGGCTGCTTCTCTCTGAAATTTGACGAGGGGTCTTTGGGCGAGACTGGCTCT 123

Db 262 TCTCCAGGGGCTGCTTCTCTCTGAAATTTGACGAGGGGTCTTTGGGCGAGACTGGCTCT 203

QY 124 GAGCGCTCCATCCCAAGGCGAGGTTCTCGTTAGCTCTGAGCGCCGACCTGGGCGCTG 183

Db 202 GAGCGCTCCATCCCAAGGCGAGGTTCTCGTTAGCTCTGAGCGCCGACCTGGGCGCTG 143

QY 184 GCGTGGATCAGGAATTTTCCAAAGAGTAGTGTCTTTGGTCTTTGGCAAACTCTAC 243

Db 142 GCGTGGATCAGGAATTTTCCAAAGAGTAGTGTCTTTGGTCTTTGGCAAACTCTAC 83

QY 244 TTAATCCAAATGGGTTTCTCTGTAAGTAGATTTTCCAAATGTAATACTTAATA 303

Db 82 TTAATCCAAATGGGTTTCTCTGTAAGTAGATTTTCCAAATGTAATACTTAATA 23

QY 304 AAGTAAAAAAAAAAAAA 325

Db 22 AAGTAAAAAAAAAAAAA 1

RESULT 12

LOCUS AI224622/c 548 bp mRNA linear EST 21-DEC-1998

DEFINITION gw56n04.x1 NCI CGAP Gas4 Homo sapiens cDNA clone IMAGE:1998967 3'

AI224622

similar to gb:J05392 SYNDSCAN-1 PRECURSOR (HUMAN);, mRNA sequence.

AI224622

EST.

AI224622.1 GI:3807335

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

REFERENCE

AUTHORS NCI-CGAP Cancer Institute, Cancer Genome Anatomy Project (CGAP),

TITLE Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapsb-r@mail.nih.gov

unknown library type

Insert Length: 2635 Std. Error: 0.00

Seq primer: -40UP from Gldco

High quality sequence stop: 352.

Location/Qualifiers

1..548

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1998967"

/clone_lib="NCI CGAP Gas4"

/tissue_type="poorly differentiated adenocarcinoma with signed ring cell features"

/lab_host="DH10B"

/note="Organ: stomach; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.69 kb. Life Technologies catalog #: 11549-011"

BASE COUNT 134 a 158 c 141 g 109 t 6 others

ORIGIN

Query Match 96.7%; Score 314.4; DB 9; Length 548;

Best Local Similarity 99.7%; Pred. No. 2.1e-41;

Matches 315; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GAGGCGCGCAGGCGCTGAGATCTCTCTGACGACGACGCCGCTGCTGCTGCGCGCG 63

Db 316 GAGGCGCGCAGGCGCTGAGATCTCTCTGACGACGACGCCGCTGCTGCTGCGCGCG 257

QY 64 TCTCCAGGGGCTGCTTCTCTCTGAAATTTGACGAGGGGTCTTTGGGCGAGACTGGCTCT 123

Db 256 TCTCCAGGGGCTGCTTCTCTCTGAAATTTGACGAGGGGTCTTTGGGCGAGACTGGCTCT 197

QY 124 GAGCGCTCCATCCCAAGGCGAGGTTCTCGTTAGCTCTGAGCGCCGACCTGGGCGCTG 183

Db 196 GAGCGCTCCATCCCAAGGCGAGGTTCTCGTTAGCTCTGAGCGCCGACCTGGGCGCTG 137

QY 184 GCGTGGATCAGGAATTTTCCAAAGAGTAGTGTCTTTGGTCTTTGGCAAACTCTAC 243

Db 136 GCGTGGATCAGGAATTTTCCAAAGAGTAGTGTCTTTGGTCTTTGGCAAACTCTAC 77

QY 244 TTAATCCAAATGGGTTTCTCTGTAAGTAGATTTTCCAAATGTAATACTTAATA 303

Db 76 TTAATCCAAATGGGTTTCTCTGTAAGTAGATTTTCCAAATGTAATACTTAATA 17

QY 304 AAGTAAAAAAAAAAAA 319

Db 16 AAGTAAAAAAAAAAAA 1

RESULT 13

LOCUS AM192078/c 519 bp mRNA linear EST 29-NOV-1999

DEFINITION x179f03.x1 NCI CGAP Pauli Homo sapiens cDNA clone IMAGE:2680925 3'

AM192078

similar to gb:J05392 SYNDSCAN-1 PRECURSOR (HUMAN);, mRNA sequence.

AM192078

EST.

AM192078.1 GI:6470777

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

REFERENCE

AUTHORS NCI-CGAP Cancer Institute, Cancer Genome Anatomy Project (CGAP),

TITLE Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapsb-r@mail.nih.gov

Life Technologies catalog #: 11548-013

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbtp/image/image.html

Seq primer: -40UP from Gldco

High quality sequence stop: 321.

Location/Qualifiers

1..519

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2680925"

/clone_lib="NCI CGAP Pauli"

/tissue_type="adenocarcinoma"

/lab_host="DH10B"

/note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.72 kb. Life Technologies catalog #: 11548-013"

BASE COUNT 127 a 153 c 131 g 108 t

ORIGIN

Query Match 96.6%; Score 314; DB 10; Length 519;

Best Local Similarity 98.4%; Pred. No. 2.6e-41;

Matches 317; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 GAGGCGCGCAGGCGCTGAGATCTCTCTGACGACGACGCCGCTGCTGCTGCGCGCG 63

Db 322 GAGGCGCGCAGGCGCTGAGATCTCTCTGACGACGACGCCGCTGCTGCTGCGCGCG 263

QY 64 TCTCCAGGGGCTGCTTCTCTCTGAAATTTGACGAGGGGTCTTTGGGCGAGACTGGCTCT 123

Db 262 TCTCCAGGGGCTGCTTCTCTCTGAAATTTGACGAGGGGTCTTTGGGCGAGACTGGCTCT 203

QY 124 GAGCGCTCCATCCCAAGGCGAGGTTCTCGTTAGCTCTGAGCGCCGACCTGGGCGCTG 183

Db 202 GAGCGCTCCATCCAGGCGAGGTTCTCCGTAGCTCTGTGTCGCCACCTGAGCCCTG 143

Qy 184 GGCTGGAATCAGAAATATTTTCCAAAGAGTGAATGCTTTTGGTTCGCAAACTAC 243

Db 142 GGTGGAAATCAGAAATATTTTCCAAAGAGTGAATGCTTTTGGTTCGCAAACTAC 83

Qy 244 TTAATCCAAATGGGTTTCTCTGTACAGTAGATTTTCCAAATGTAATTAATTAATA 303

Db 82 TTAATCCAAATGGGTTTCTCTGTACAGTAGATTTTCCAAATGTAATTAATTAATA 23

Qy 304 AAGTAAAAAATTTTAAAAA 325

Db 22 AAGTAAAAAATTTTAAAAA 1

RESULT 14

LOCUS AM662594 461 bp mRNA linear EST 06-APR-2000

DEFINITION h133a10.x1 NCI CGAP Col4 Homo sapiens cDNA clone IMAGE:2974074 3', similar to gb:J05392 SYNDSCAN-1 PRECURSOR (HUMAN);, mRNA sequence.

ACCESSION AM662594

VERSION AM662594.1 GI:7455133

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Washington University Genome Sequencing Center
found through the I.M.A.G.E. Consortium/HLN at: image.llnl.gov/image/html/resources.shtml

FEATURES

source

1. 461

Location/Qualifiers

/db_xref="taxon:9606"

/clone="IMAGE:2974074"

/clone_lib="NCI-CGAP Col4"

/tissue_type="moderately-differentiated adenocarcinoma"

/lab_host="DH10B"

/note="Organ: colon; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.7 Kb. Life Technologies catalog #: 11531-019"

BASE COUNT 112 a 123 c 118 g 108 t

ORIGIN

Query Match 96.5%; Score 313.6; DB 10; Length 461;
Best Local Similarity 98.8%; Pred. No. 3.2e-41;
Matches 316; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 AAGGCGGAGGAGGCTGAGATCTCTGAGACACAGCCGCTGCTGAGGCGCGT 64

Db 320 AAGGCGGAGGAGGCTGAGATCTCTGAGACACAGCCGCTGCTGAGGCGCGT 261

Qy 65 CTCGAGGAGGCTCTCTCTCTGGAATTAACAGAGGAGTCTTGGCAGAGCTGCTTG 124

Db 260 CTCGAGGAGGCTCTCTCTCTGGAATTAACAGAGGAGTCTTGGCAGAGCTGCTTG 201

Qy 125 AAGCGCTCCATCCAGGCGAGGTTCTCCGTAGCTCTGTGTCGCCACCTGAGCCCTG 184

Db 200 AAGCGCTCCATCCAGGCGAGGTTCTCCGTAGCTCTGTGTCGCCACCTGAGCCCTG 141

Qy 185 GCTGGAATCAGAAATATTTTCCAAAGAGTGAATGCTTTTGGTTCGCAAACTACT 244

Db 140 GCTGGAATCAGAAATATTTTCCAAAGAGTGAATGCTTTTGGTTCGCAAACTACT 81

Qy 245 TTAATCCAAATGGGTTTCTCTGTACAGTAGATTTTCCAAATGTAATTAATTAATA 304

Db 80 TTAATCCAAATGGGTTTCTCTGTACAGTAGATTTTCCAAATGTAATTAATTAATA 21

Qy 305 AAGTAAAAAATTTTAAAAA 324

Db 20 AAGTAAAAAATTTTAAAAA 1

RESULT 15

LOCUS A1691025 636 bp mRNA linear EST 14-DEC-1999

DEFINITION tg13e11.x1 NCI-CGAP Utr3 Homo sapiens cDNA clone IMAGE:2208716 3', similar to gb:J05392 SYNDSCAN-1 PRECURSOR (HUMAN); contains element PRR5 repetitive element ;, mRNA sequence.

ACCESSION A1691025

VERSION A1691025.1 GI:4902327

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Washington University Genome Sequencing Center
found through the I.M.A.G.E. Consortium/HLN at: www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 1849 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 337.

FEATURES

source

1. 636

Location/Qualifiers

/db_xref="taxon:9606"

/clone="IMAGE:2208716"

/clone_lib="NCI-CGAP Utr3"

/tissue_type="poorly-differentiated endometrial adenocarcinoma, 2 pooled tumors"

/lab_host="DH10B"

/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.45 Kb. Life Technologies catalog #: 11541-018"

BASE COUNT 157 a 173 c 152 g 154 t

ORIGIN

Query Match 96.1%; Score 312.4; DB 9; Length 636;
Best Local Similarity 98.1%; Pred. No. 4.1e-41;
Matches 316; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 GAGGCGGAGGAGGCTGAGATCTCTGAGACACAGCCGCTGCTGAGGCGCGG 63

Db 345 GAGGCGGAGGAGGCTGAGATCTCTGAGACACAGCCGCTGCTGAGGCGCGG 286

Qy 64 TCTCAGGAGGCTCTCTCTCTGGAATTAACAGAGGAGTCTTGGGCAAGAGCTGCTCT 123

```

Db 285 TCTCCAGGGGCTCTTCTCCTGGAATTGACGAGGGGTGTCTTGCGAAAGCTGGCTTT 226
QY 124 GAGCGCTTCATCCAGGCCAGGTTCTCCGTAGCTCTGTGGCCCCACCCCTGGGCTTG 183
Db 225 GAGGCGCTCCATCCAGGCCAGGTTCTCCGTAGCTCTGTGGCCCCCTGGGCTTG 166
QY 184 GGCTGGAATCAGGAATATTTTCCAAAGAGTATGCTTTTGGCTTTGGCAAACTCTAC 243
Db 165 GGCTGGAATCAGGAATATTTTCCAAAGAGTATGCTTTTGGCAAACTCTAC 106
QY 244 TTAATCCAAATGGGTTTCTCTGTACGTAGATTTTCCAAATGTATATACTTAATATA 303
Db 105 TTAATCCAAATGGGTTTCTCTGTACGTAGATTTTCCAAATGTATATACTTAATATA 46
QY 304 AAGTAAAAAAAAAAAAAAAAA 325
Db 45 AAGTAAAAAAAAAAAAAAAAA 24

```

Search completed: January 10, 2003, 03:20:53
 Job time : 701.746 secs